

Article

Sensor Selection for Decentralized Large-Scale Multi-Target Tracking Network

Feng Lian *, Liming Hou, Bo Wei and Chongzhao Han

Ministry of Education Key Laboratory for Intelligent Networks and Network Security (MOE KLINNS), School of Electronics and Information Engineering, Xi'an Jiaotong University, Xi'an 710049, China; hliming2017@stu.xjtu.edu.cn (L.H.); weibo18291927630@163.com (B.W.); czhan@xjtu.edu.cn (C.H.);

* Correspondence: lianfeng1981@xjtu.edu.cn; Tel.: +86-158-2907-9318

Received: 11 October 2018; Accepted: 20 November 2018; Published: 23 November 2018



Abstract: A new optimization algorithm of sensor selection is proposed in this paper for decentralized large-scale multi-target tracking (MTT) network within a labeled random finite set (RFS) framework. The method is performed based on a marginalized δ -generalized labeled multi-Bernoulli RFS. The rule of weighted Kullback-Leibler average (KLA) is used to fuse local multi-target densities. A new metric, named as the label assignment (LA) metric, is proposed to measure the distance for two labeled sets. The lower bound of LA metric based mean square error between the labeled multi-target state set and its estimate is taken as the optimized objective function of sensor selection. The proposed bound is obtained by the information inequality to RFS measurement. Then, we present the sequential Monte Carlo and Gaussian mixture implementations for the bound. Another advantage of the bound is that it provides a basis for setting the weights of KLA. The coordinate descent method is proposed to compromise the computational cost of sensor selection and the accuracy of MTT. Simulations verify the effectiveness of our method under different signal-to-noise ratio scenarios.

Keywords: sensor selection; multi-target tracking; labeled random finite set; decentralized sensor network; error bound

1. Introduction

With the development of communication and information fusion technologies, multi-target tracking (MTT) [1] based on sensor network becomes a new research hotspot. In general, the sensor networks are divided into two main categories according to their structure: one is centralized network and the other is decentralized network. Compared with the centralized network, the decentralized network is more widely concerned because of its parallelism, flexibility, robustness, scalability, anti-interference and fault tolerance, et al. In most practical applications, due to the limitations of communication bandwidth, energy consumption, computational cost, storage space et al, not all sensors in a network can be activated to observe targets at the same instant. As a result, the problem of sensor selection arises from this, which actually belongs to a branch of sensor management [2]. Although some research results [3–8] have been proposed for it, none of them jointly consider the uncertainty of target number and data association.

In the past two decades, random finite set (RFS) [9] based MTT has attracted extensive attention. By the use of RFS, MTT is described as a Bayesian estimation of state and observation sets. RFS filtering has been developed from the primal probability hypothesis density (PHD) [10–12], cardinalized PHD [13,14] and multi-Bernoulli [15,16] filters to the latest δ -generalized labeled multi-Bernoulli (δ -GLMB) filter [17–19]. The advantage of the latter is its conjugacy and track formation. Nevertheless, the number of components involved in the GLMB density increases exponentially with the recursion. Therefore, two approximation methods, the LMB and Marginalized δ -GLMB (M δ -GLMB) [20–22], are

subsequently proposed to reduce the computational cost of GLMB. They are also more suitable for multi-sensor scenarios. [23] has shown that the filtering accuracy of $M\delta$ -GLMB is close to δ -GLMB and the two are significantly superior to LMB in the scenarios of low signal-to-noise ratio (SNR).

Besides the δ -GLMB conjugate prior, the Poisson multi-Bernoulli mixture (PMBM) filter [24,25] and multi-Bernoulli mixture (MBM) filter [25] are also conjugate priors. Track formation in the (P)MBM formulation can also be attained using RFS of trajectories [26].

In recent years, although the RFS-based methods have been used to control the position of one or several mobile sensors for MTT [27–34], none of them refer to the sensor network. Actually, in many cases, the sensors in a network are immobile. Instead, the problems of structure, constraints, node selection and so forth become rather important especially for a large-scale sensor network.

As a result, this paper focuses on the emerging problem of sensor selection for decentralized large-scale MTT networks. A new optimization algorithm for sensor selection is proposed based on the $M\delta$ -GLMB filter. In the proposed method, the fusion of local multi-target posterior densities is carried out by using the rule of weighted Kullback-Leibler average (KLA) [23].

The main contributions of our method includes four aspects. First, the sensor selection is described as a constrained optimization problem with a Bayesian recursion of labeled multi-target RFS. A new metric, named as the label assignment (LA) metric, is proposed to measure the distance for two labeled sets. The lower bound of LA metric based mean square error (MSE) between the labeled multi-target state set and its estimate is treated as the optimized objective function of sensor selection. The bound is derived by the information inequality to RFS measurement [35]. The detailed proofs for the LA metric and its lower bound are presented in the appendices. Second, the normalized weights of the KLA rule are set according to the proposed bound. Third, both the sequential Monte Carlo (SMC) [10,36] and Gaussian mixture (GM) [11–16] implementations for the bound are presented. Fourth, because the computational cost of selection optimization increases with sensor number in the form of combination explosion, a sub-optimization method called coordinate descent [37] is proposed to compromise the computational cost and tracking accuracy.

The simulation results show that when the sensors in a decentralized large-scale network have different observation performance, 1) the MTT accuracy of our method is much better than that of the Cauchy-Schwarz (CS) divergence based methods [31,32]; 2) compared with the genetic algorithm [38], the coordinate descent method significantly shortens the calculation time of sensor selection; 3) the GM implementation of the bound is obviously faster than its SMC implementation.

2. Mathematical Background

2.1. Labeled RFS and $M\delta$ -GLMB

In this paper, the unlabeled and labeled variables are, respectively, represented by the italics and bold. For example, the unlabeled state, measurement and their sets are noted as x , z , X and Z ; the labeled state and its set are noted as $\mathbf{x} = (x, \ell)$ and \mathbf{X} , where ℓ is the discrete label of \mathbf{x} . Let $\mathcal{L}(\mathbf{X})$, $|\mathbf{X}|$ and $\mathbb{X} \times \mathbb{L}$ denote the label set, cardinality and space of \mathbf{X} , where \mathbb{X} and \mathbb{L} are the spaces of the unlabeled state and label.

The state estimates of single-target and multi-targets derived from a measurement set Z are both the functions of Z . To make this clearer, they are, respectively, noted as $\hat{x}(Z)$ and $\hat{X}(Z)$. $\hat{\mathbf{x}}(Z)$ and $\hat{\mathbf{X}}(Z)$ are their labeled versions.

Let $\delta_Y(X)$, $1_Y(X)$ and $p^{\mathbf{X}}$ denote the functions of generalized Kronecker, inclusion indicator and multi-object exponential,

$$\delta_Y(X) = \begin{cases} 1, & \text{if } X = Y \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

$$1_Y(X) = \begin{cases} 1, & \text{if } X \subseteq Y \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

$$p^{\mathbf{X}} = \begin{cases} \prod_{\mathbf{x} \in \mathbf{X}} p(\mathbf{x}), & \mathbf{X} \neq \emptyset \\ 1, & \mathbf{X} = \emptyset \end{cases} \tag{3}$$

where $1_Y(\{x\})$ is abbreviated as $1_Y(x)$. Furthermore, if $1_Y(X) = 1$, then let $Y - X$ denote the complementary set of X in Y . $Y - \{x\}$ is abbreviated as $Y - x$.

For any real-valued function $b(\mathbf{X})$ of \mathbf{X} , its set integral $\int b(\mathbf{X})\delta\mathbf{X}$ is defined as

$$\int b(\mathbf{X})\delta\mathbf{X} = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{\ell_{1:n} \in \mathbb{L}_n} \int_{\mathbb{X}_n} b(\mathbf{X}_n) dx_{1:n} \tag{4}$$

where $x_{1:n} = x_1, \dots, x_n$ and $\ell_{1:n} = \ell_1, \dots, \ell_n$, $\mathbf{X}_n = \{x_{1:n}\}$ is a n -element labeled set, \mathbb{X}_n and \mathbb{L}_n are the spaces of X_n and $\ell_{1:n}$.

If \mathbf{X} is a $M\delta$ -GLMB RFS, then its density is described as [22]

$$\pi(\mathbf{X}) = \Delta(\mathbf{X}) \sum_{I \in \mathcal{F}(\mathbb{L})} \delta_I(\mathcal{L}(\mathbf{X})) \omega_I p_I^{\mathbf{X}} \tag{5}$$

where $\Delta(\mathbf{X}) = \delta_{|\mathbf{X}|}(|\mathcal{L}(\mathbf{X})|)$ is a distinct indicator for the labels of \mathbf{X} , $I \in \mathcal{F}(\mathbb{L})$ is a label set in the collection $\mathcal{F}(\mathbb{L})$ of finite subsets of \mathbb{L} , the weight ω_I is the existing probability of the label set I , $p_I(\mathbf{x})$ is the density of \mathbf{x} involved in I . The $M\delta$ -GLMB density is abbreviated as $\pi = \{(\omega_I, p_I)\}_{I \in \mathcal{F}(\mathbb{L})}$ and its cardinality distribution is

$$P(|\mathbf{X}| = n) = \sum_{I \in \mathcal{F}_n(\mathbb{L})} \omega_I \tag{6}$$

where $\mathcal{F}_n(\mathbb{L})$ is the collection of n -element subsets of \mathbb{L} .

2.2. Information Inequality to RFS Measurement

Let $\hat{x}(Z_m)$ be an unbiased estimate of x derived from an m -element measurement set Z_m and $f(x, Z_m)$ be a joint density over the space $\mathbb{X}_1 \times \mathbb{Z}_m$. Assuming that regularity conditions hold and $\partial^2 \log f(x, Z_m) / \partial x^i \partial x^j$ exists, the information inequality to RFS measurement is [35]

$$\int_{\mathbb{Z}_m} \int_{\mathbb{X}_1} f(x, Z_m) \left(x^l - \hat{x}^l(Z_m)\right)^2 dx dz_{1:m} \geq [J_m^{-1}]^{l,l}, \quad l = 1, \dots, L \tag{7}$$

where $z_{1:m} = z_1, \dots, z_m$, L is the dimension of x , x^l and $\hat{x}^l(Z_m)$ are the l th components of the vectors x and $\hat{x}(Z_m)$, J_m is the $L \times L$ Fisher information matrix (FIM) given $|Z| = m$,

$$[J_m]^{i,j} = -E_f \left[\frac{\partial^2 \log f(x, Z_m)}{\partial x^i \partial x^j} \right] = - \int_{\mathbb{Z}_m} \int_{\mathbb{X}_1} f(x, Z_m) \frac{\partial^2 \log f(x, Z_m)}{\partial x^i \partial x^j} dx dz_{1:m}, \quad i, j = 1, \dots, L \tag{8}$$

(7) holds with equality if and only if $f(x, Z_m)$ satisfies the distribution of exponential family.

2.3. A New Metric for Labeled RFS

It is well known that the optimal sub-pattern assignment (OSPA) metric [39–41] has been extensively used to measure the distance for two unlabeled sets. Although the OSPA metric could also measure differences in the set labels, it may not be very appropriate for the labeled RFS in some application scenarios. As a result, a new metric between two labeled sets \mathbf{X} and \mathbf{Y} of order $1 \leq p \leq \infty$ with cut-off $c > 0$ is proposed as follows.

$$\bar{d}_p^{(c)}(\mathbf{X}, \mathbf{Y}) = \begin{cases} 0 & |\mathbf{X}| = |\mathbf{Y}| = 0 \\ \left(\frac{\sum_{\ell \in \mathcal{L}(\mathbf{X}) \cap \mathcal{L}(\mathbf{Y})} d^{(c)}(x_\ell, y_\ell)^p + c^p (|\mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Y})| - |\mathcal{L}(\mathbf{X}) \cap \mathcal{L}(\mathbf{Y})|)}{|\mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Y})|} \right)^{\frac{1}{p}} & |\mathbf{X}| + |\mathbf{Y}| > 0 \end{cases} \tag{9}$$

where $\mathcal{L}(\mathbf{X})$ and $\mathcal{L}(\mathbf{Y})$ are the label sets of \mathbf{X} and \mathbf{Y} , $x_\ell \in X$ and $y_\ell \in Y$ are the unlabeled elements corresponding to the label ℓ , $|\mathcal{L}(\mathbf{X}) \cap \mathcal{L}(\mathbf{Y})|$ and $|\mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Y})| - |\mathcal{L}(\mathbf{X}) \cap \mathcal{L}(\mathbf{Y})|$ respectively indicates the number of elements in \mathbf{X} and \mathbf{Y} which have the same and different labels,

$$d^{(c)}(x_\ell, y_\ell) = \min(c, \|x_\ell - y_\ell\|) \quad (10)$$

denotes the 2-norm of x_ℓ and y_ℓ cut off at $c > 0$.

See Appendix A for the proof that the $\bar{d}_p^{(c)}(\mathbf{X}, \mathbf{Y})$ is a metric.

The proposed metric, which is named as the LA metric by us, is a different metric than the OSPA. A physical meaning about the LA metric between two labeled sets \mathbf{X} and \mathbf{Y} is explained as follows. For all $\mathbf{x} \in \mathbf{X}$ and $\mathbf{y} \in \mathbf{Y}$, if \mathbf{x} has the same label as \mathbf{y} , then \mathbf{x} is paired with \mathbf{y} and a 'location' error $d^{(c)}(x, y)$ between them is involved in the metric; otherwise, \mathbf{x} is unpaired with \mathbf{y} and a 'penalty' error c for them is involved in the metric.

The most significant difference between the OSPA metric and the LA metric is: In the OSPA metric, the elements of \mathbf{X} are paired with the elements of \mathbf{Y} depending on the optimal assignment distance of their unlabeled versions. In contrast, in the LA metric, the elements of \mathbf{X} are paired with the elements of \mathbf{Y} completely depending on their labels.

Take the MTT with the labeled RFS state for example. The calculation of the OSPA error may pair an estimate with a state due to the rule of optimal assignment even if they have different labels. Instead, the calculation of LA error prohibits this kind of pairing even if their unlabeled states are sufficiently close to each other. As a result, the LA metric involves not only the estimation error arising from the target number and individual states as the OSPA metric but also the additional estimation error arising from the labels. In this sense, the LA metric is more demanding than the OSPA metric for measuring the error between the labeled state set and its estimate.

3. Problem Formulation

To simplify the formulas, the time index is omitted and the subscript '+' is used to indicate the predicted density.

Multiple targets independently move in region \mathcal{A} with random birth and death. The multi-target states are modeled as a labeled RFS \mathbf{X} . The dynamic of a single state $\mathbf{x} = (x, \ell) \in \mathbf{X}$ is described by the survival probability $p_s(\mathbf{x})$ and transition density $f(\mathbf{x}|\mathbf{x}')\delta_{\ell'}(\ell)$. The dynamic of multi-target states is described by the transition density $f(\mathbf{X}|\mathbf{X}')$. Here $\mathbf{x}' = (x', \ell')$ and \mathbf{X}' are, respectively, the state and state set at the last time.

Targets are observed by the decentralized sensor network shown in Figure 1. The network is composed of sensor nodes (SNs) and local fusion centers (LFCs). Each SN receives measurements and communicates with its superior LFC. Each LFC receives the measurements, conducts data processing and storage, communicates with the other LFCs connected to it and manages its subordinate SNs.

The network structure is completely described by a topological graph with parameter $\{N, C, A\}$, where N and C are the label sets of SN and LFC, $A \subseteq C \times C$ is the set of directed connections between LFCs. If the LFC j can receive data of the LFC i , then $(i, j) \in A$. Let $C^j = \{i \in C : (i, j) \in A\}$ be the label set of the LFCs connected to the LFC j (including itself) and N^j be the label set of the SNs belonging to the LFC j . Each SN only belongs to one LFC, which indicates $\bigcap_{j \in C} N^j = \emptyset$ and $\bigcup_{j \in C} N^j = N$.

The most significant difference between the decentralized network and the centralized or hierarchical network is that the former has no global fusion center connected to all SNs or all LFCs. The network structure remains unchanged and all measurements are synchronized during the monitoring period.

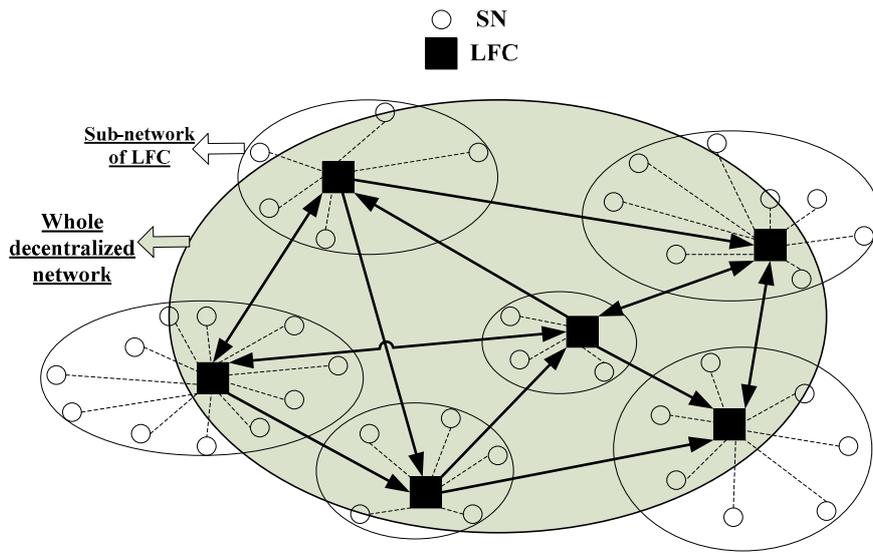


Figure 1. Diagram for decentralized sensor network. \circ and \blacksquare denote SN and LFC, solid lines with arrows denote directed connections between LFCs, the dotted lines denote connections between LFC and its subordinate SNs.

For the SN $s \in N$, it may receive clutter and target measurement or miss the detection. The measurements are modeled as a RFS Z^s over the space \mathbb{Z}^s . $z^s \in Z^s$ is a single measurement. Clutter is modeled as a Poisson RFS with intensity $\kappa^s(z^s)$, and

$$\lambda^s = \int_{\mathbb{Z}_1^s} \kappa^s(z^s) dz^s \tag{11}$$

is the clutter rate.

The multi-target likelihood of the SN s is obtained from [9] as

$$g^s(Z^s | \mathbf{X}) = e^{-\lambda^s} [\kappa^s]^{Z^s} \sum_{\theta^s \in \Theta^s} [\psi_{Z^s}^s(\cdot; \theta^s)]^{\mathbf{X}} \tag{12}$$

where

$$\psi_{Z^s}^s(\mathbf{x}; \theta^s) = \delta_0(\theta^s(\ell))(1 - p_d^s(\mathbf{x})) + (1 - \delta_0(\theta^s(\ell))) \frac{p_d^s(\mathbf{x}) g^s(z_{\theta^s(\ell)}^s | \mathbf{x})}{\kappa^s(z_{\theta^s(\ell)}^s)} \tag{13}$$

where $p_d^s(\mathbf{x})$ and $g^s(z^s | \mathbf{x})$ are the single-target detection probability and likelihood, Θ^s is a collection of association mapping $\theta^s : \mathcal{L}(\mathbf{X}) \rightarrow \{0, 1, \dots, |Z^s|\}$. $\theta^s(\ell) > 0$ or $\theta^s(\ell) = 0$ indicates that the track $\ell \in \mathcal{L}(\mathbf{X})$ generates a measurement $z_{\theta^s(\ell)}^s \in Z^s$ or be missed. Each track at most generates one measurement and each measurement is at most generated by one track, which indicates that $\ell = \ell'$ if $\theta^s(\ell) = \theta^s(\ell') > 0$. The number of association hypotheses is [42]

$$\chi_{|\mathcal{L}(\mathbf{X})|, |Z^s|} = \sum_{i=0}^{\min(|\mathcal{L}(\mathbf{X})|, |Z^s|)} \frac{|\mathcal{L}(\mathbf{X})|! \cdot |Z^s|!}{(|\mathcal{L}(\mathbf{X})| - i)! \cdot (|Z^s| - i)! \cdot i!} \tag{14}$$

Due to some restrictions, only part of SNs involved in the sub-network of each LFC can be activated to observe targets at each scan. Assume that the multi-target likelihoods of all the SNs in the network are independent of each other given the labeled state set \mathbf{X} . Algorithm 1 presents the steps of sensor selection and MTT for the LFC $j \in C$ under a Bayesian framework. Note that the sequence of measurement sets up to the last time is omitted here for simplifying the formulas of Bayesian recursion.

Algorithm 1. Sensor selection and MTT for the LFC j .

1. **Prediction:** Calculate the current predicted density by $\pi_+^j(\mathbf{X}) = \int f(\mathbf{X}|\mathbf{X}')\bar{\pi}^j(\mathbf{X}')\delta\mathbf{X}'$, where $\bar{\pi}^j(\mathbf{X}')$ is the fused density at the last time;
 2. **SN selection:** Select the SN subset $S^j \subseteq N^j$ and receive a collection of their measurement sets Z^{S^j} ;
 3. **Update:** Calculate the current posterior density by $\pi^j(\mathbf{X}|Z^{S^j}) = \frac{\prod_{s \in S^j} g^s(Z^s|\mathbf{X})\pi_+^j(\mathbf{X})}{\int \prod_{s' \in S^j} g^{s'}(Z^{s'}|\mathbf{X})\pi_+^j(\mathbf{X})\delta\mathbf{X}}$ and then transmit the density to the LFCs connected to the LFC j ;
 4. **Fusion:** Receive the posterior densities from the LFC set C^j and then calculate the fused density by the weighted KLA rule $\bar{\pi}^j(\mathbf{X}) = \frac{\prod_{i \in C^j} [\pi^i(\mathbf{X}|Z^{S^i})]^{a^{i,j}}}{\int \prod_{i' \in C^j} [\pi^{i'}(\mathbf{X}|Z^{S^{i'}})]^{a^{i',j}} \delta\mathbf{X}}$, where $a^{i,j}$ ($i \in C^j$) is the preset normalized weight;
 5. **State extraction:** Extract the current state estimate from $\bar{\pi}^j(\mathbf{X})$ as the output. Go to Step 1.
-

Remark 1. It can be seen from Steps 3 and 4 of Algorithm 1 that in this network, each LFC only communicates once with the LFCs connected to it per recursion. Because of this, the consensus iteration [43,44] of each LFC cannot be carried out since each LFC has to communicate with other LFCs more than once in the iterative process. Finally, the consensus fusion [23] over the entire decentralized sensor network cannot be achieved. As a result, the multi-target estimates outputted by different LFCs may be different. In fact, this character is consistent with most practical application systems.

The SN selection in Step 2 of Algorithm 1 is described as the following constrained optimization problems:

$$\begin{aligned}
 [S^j]^* &= \underset{S^j \subseteq N^j}{\operatorname{argmin}/\operatorname{max}} \vartheta^j(S^j; \pi_+^j) \\
 \text{s.t.} \quad &\begin{cases} \gamma_i^j(S^j; \pi_+^j) \geq 0 & i = 1, \dots, l \\ v_k^j(S^j; \pi_+^j) = 0 & k = 1, \dots, m \end{cases}
 \end{aligned} \tag{15}$$

where $\vartheta^j(S^j; \pi_+^j)$, $\gamma_i^j(S^j; \pi_+^j) \geq 0$ and $v_k^j(S^j; \pi_+^j) = 0$ are the objective function, inequality and equality constraints of S^j given the predicted density $\pi_+^j(\mathbf{X})$.

The ultimate goal for a MTT network is to optimize the tracking precision. As is known to all, the MSE between target state and its estimate is currently the most widely-used evaluation indicator for tracking accuracy. Given the selected SN subset S^j for the LFC $j \in C$, the MSE $[\sigma_{S^j}^j]^2$ between \mathbf{X} and its Bayesian estimate $\hat{\mathbf{X}}^j(Z^{S^j})$ is

$$\begin{aligned}
 [\sigma_{S^j}^j]^2 &= E \left[e^2(\mathbf{X}, \hat{\mathbf{X}}^j(Z^{S^j})) \right] \\
 &= \int_{\mathbb{Z}^{S^j}} \int_{\mathbb{X} \times \mathbb{L}} f(\mathbf{X}, Z^{S^j}) e^2(\mathbf{X}, \hat{\mathbf{X}}^j(Z^{S^j})) \delta\mathbf{X} \delta Z^{S^j} \\
 &= \int_{\mathbb{Z}^{S^j}} \int_{\mathbb{X} \times \mathbb{L}} \prod_{s \in S^j} g^s(Z^s|\mathbf{X}) \pi_+^j(\mathbf{X}) e^2(\mathbf{X}, \hat{\mathbf{X}}^j(Z^{S^j})) \delta\mathbf{X} \delta Z^{S^j}
 \end{aligned} \tag{16}$$

where \mathbb{Z}^{S^j} is the joint measurement space of the SN set S^j , $f(\mathbf{X}, Z^{S^j})$ is the joint density of (\mathbf{X}, Z^{S^j}) , $e(\mathbf{X}, \hat{\mathbf{X}}^j(Z^{S^j}))$ denotes the error distance between \mathbf{X} and $\hat{\mathbf{X}}^j(Z^{S^j})$. In this paper, $e(\mathbf{X}, \hat{\mathbf{X}}^j(Z^{S^j}))$ is measured by the 2nd-order LA metric $\bar{d}_2^{(c)}(\mathbf{X}, \hat{\mathbf{X}}^j(Z^{S^j}))$ in (9).

Nevertheless, $[\sigma_{S^j}^j]^2$ cannot be used as the objective function of sensor selection optimization in (15) because $\hat{\mathbf{X}}^j(Z^{S^j})$ is unknown before sensor selection. To solve this, we replace $[\sigma_{S^j}^j]^2$ with its lower bound $[\underline{\sigma}_{S^j}^j]^2$, which provides an online indication on the limit of MTT accuracy within the labeled RFS framework.

Treating $\pi_+^j(\mathbf{X})$ as a default condition, (15) is finally rewritten as

$$[S^j]^* = \underset{S^j \subseteq N^j}{\operatorname{argmin}} [\underline{\sigma}_{S^j}^j]^2$$

$$\text{s.t.} \begin{cases} \gamma_i^j(S^j) \geq 0 & i = 1, \dots, l \\ \nu_k^j(S^j) = 0 & k = 1, \dots, m \end{cases} \quad (17)$$

4. Lower Bound For LA Metric Based MSE and Sub-Optimization For Sensor Selection

4.1. Derivation of LA Bound

In Section 4.1, Section 4.2, Section 4.3 and Appendix B, the superscript ‘j’ for the index of LFC is omitted. For example, $(\underline{\sigma}_{S^j}^j)^2$ is abbreviated as $\underline{\sigma}_S^2$.

In order to derive $\underline{\sigma}_S^2$, it is assumed that

A1: Multi-target Bayesian recursion is a M δ -GLMB RFS [22]. As a result, the predicted density $\pi_+(\mathbf{X})$ and posterior density $\pi(\mathbf{X}|Z^S)$ can be described as $\pi_+ = \{(\omega_{I,+}, p_{I,+})\}_{I \in \mathcal{F}(\mathbb{L})}$ and $\pi(\cdot|Z^S) = \{(\omega_I(Z^S), p_I(\cdot|Z^S))\}_{I \in \mathcal{F}(\mathbb{L})}$.

A2: Although the optimal estimate of \mathbf{X} can be extracted from the fused density $\bar{\pi}(\mathbf{X})$ by using the joint or marginal multi-target estimator [9], both the methods are very difficult to be implemented. Alternatively, the target number is firstly estimated according to the maximum a posterior (MAP) criterion and then the individual states are estimated according to the unbiased criterion under the obtained target number. In fact, the suboptimal method is applied in almost all multi-target Bayesian filters.

Let $Z_{m^S}^S = Z_{m^{s_1}}^{s_1}, \dots, Z_{m^{s_{|S|}}}^{s_{|S|}}$ be the collection of measurement sets from the SN set S and $\mathbb{Z}_{m^S}^S = \mathbb{Z}_{m^{s_1}}^{s_1} \times \dots \times \mathbb{Z}_{m^{s_{|S|}}}^{s_{|S|}}$ be the space of $Z_{m^S}^S$, where m^{s_i} is the number of measurements received by the SN s_i . Let $q(\mathbf{X}_n, Z_{m^S}^S)$ be the joint density over the space $(\mathbb{X} \times \mathbb{L})_n \times \mathbb{Z}_{m^S}^S$. According to Bayesian formula, $q(\mathbf{X}_n, Z_{m^S}^S)$ is written as

$$q(\mathbf{X}_n, Z_{m^S}^S) = \frac{1}{\Omega_{n,m^S}} \prod_{s \in S} g^s(Z_{m^s}^S | \mathbf{X}_n) \pi_+(\mathbf{X}_n) \quad (18)$$

where Ω_{n,m^S} is a normalization factor,

$$\Omega_{n,m^S} = \sum_{\ell_{1:n} \in \mathbb{L}_n} \int_{\mathbb{Z}_{m^S}^S} \int_{\mathbb{X}_n} \prod_{s \in S} g^s(Z_{m^s}^S | \mathbf{X}_n) \pi_+(\mathbf{X}_n) dx_{1:n} dz_{1:m^S}^S \quad (19)$$

where $\int_{\mathbb{Z}_{m^S}^S} \cdot dz_{1:m^S}^S = \int_{\mathbb{Z}_{m^{s_{|S|}}}^{s_{|S|}}} \dots \int_{\mathbb{Z}_{m^{s_1}}^{s_1}} \cdot dz_{1:m^{s_1}}^{s_1} \dots dz_{1:m^{s_{|S|}}}^{s_{|S|}}$. (19) shows that $\Omega_{n,m^S} / (m^S! \cdot n!)$ is actually the probability $P(|\mathbf{X}| = n, |Z^S| = m^S)$, where $m^S! = m^{s_1}! \dots m^{s_{|S|}}!$ and $|Z^S| = m^S$ denotes $|Z^{s_1}| = m^{s_1}, \dots, |Z^{s_{|S|}}| = m^{s_{|S|}}$. Substituting **A1** and (12) into (19) as well as using Lemma 12 in [17], Ω_{n,m^S} is obtained as

$$\Omega_{n,m^S} = n! \left(\prod_{s \in S} e^{-\lambda^s} [\lambda^s]^{m^s} \right) \sum_{I \in \mathcal{F}_n(\mathbb{L})} \omega_{I,+} \sum_{\theta^S \in \Theta^S} \varphi_I(\theta^S) \quad (20)$$

where $\theta^S \in \Theta^S$ denotes $\theta^{s_1} \in \Theta^{s_1}, \dots, \theta^{s_{|S|}} \in \Theta^{s_{|S|}}$,

$$\varphi_I(\theta^S) = \left\langle p_{I,+}(\cdot, \ell), \prod_{s \in S} \left(\delta_0(\theta^s(\ell)) (1 - p_d^s(\cdot, \ell)) + (1 - \delta_0(\theta^s(\ell))) \frac{p_d^s(\cdot, \ell)}{\lambda^s} \right) \right\rangle^I \quad (21)$$

Assume that $\theta^s(\ell) > 0$ if $s \in Y$ and $\theta^s(\ell) = 0$ if $s \in (S - Y)$. Then, (21) can be rewritten as

$$\varphi_I(\theta^S) = \left(\sum_{Y \subseteq S} \frac{\langle p_{I,+}(\cdot, \ell), p_d^{Y,S}(\cdot, \ell) \rangle}{\prod_{s \in Y} \lambda^s} \right)^I = \varphi_I(S) \tag{22}$$

where $\langle p_{I,+}(\cdot, \ell), p_d^{Y,S}(\cdot, \ell) \rangle = \int_{\mathbb{X}_1} p_{I,+}(x, \ell) p_d^{Y,S}(x, \ell) dx$ denotes the inner product corresponding to x ,

$$p_d^{Y,S}(\mathbf{x}) = \prod_{s \in Y} p_d^s(\mathbf{x}) \prod_{s' \in (S-Y)} (1 - p_d^{s'}(\mathbf{x})) \tag{23}$$

denotes the probability that only the SN subset $Y \subseteq S$ receives the measurement from state \mathbf{x} while the others miss the measurement.

(22) shows that $\varphi_I(\theta^S)$ is independent of the association mapping θ^S . From (14), (20) and (22), Ω_{n,m^S} is finally obtained as

$$\Omega_{n,m^S} = n! \left(\prod_{s \in S} e^{-\lambda^s} [\lambda^s]^{m^s} \chi_{n,m^s} \right) \sum_{I \in \mathcal{F}_n(\mathbb{L})} \omega_{I,+} \varphi_I(S) \tag{24}$$

Since $q(\mathbf{X}_n, Z_{m^S}^S)$ is permutation invariant over $\mathbf{x}_{1:n}$, its marginal density over any of $\mathbf{x}_{1:n}$ is the same and is obtained by

$$q_n(\mathbf{x}, Z_{m^S}^S) = \int_{\mathbb{X}_{n-1}} q(\{\mathbf{x}, \mathbf{x}_{2:n}\}, Z_{m^S}^S) d\mathbf{x}_{2:n} \tag{25}$$

Substituting (18) into (25) as well as using **A1** and the identical equation $\delta_n(|\{\ell, \ell_{2:n}\}|) = \delta_{n-1}(|\{\ell_{2:n}\}|) (1 - 1_{\{\ell_{2:n}\}}(\ell))$, $q_n(\mathbf{x}, Z_{m^S}^S)$ is written as

$$q_n(\mathbf{x}, Z_{m^S}^S) = \frac{1}{\Omega_{n,m^S}} \sum_{\ell_{2:n} \in \mathbb{L}_{n-1}} \delta_{n-1}(|\{\ell_{2:n}\}|) (1 - 1_{\{\ell_{2:n}\}}(\ell)) \sum_{I \in \mathcal{F}_n(\mathbb{L})} \omega_{I,+} \delta_I(\{\ell, \ell_{2:n}\}) \cdot \int_{\mathbb{X}_{n-1}} \prod_{s \in S} g^s(Z_{m^S}^S | \{\mathbf{x}, \mathbf{x}_{2:n}\}) p_{I,+}(\mathbf{x}) \prod_{t=2}^n p_{I,+}(\mathbf{x}_t) d\mathbf{x}_{2:n} \tag{26}$$

Substituting (12) into (26) and then simplifying the result, we get

$$q_n(\mathbf{x}, Z_{m^S}^S) = \frac{1}{\Omega_{n,m^S}} \left(\prod_{s \in S} e^{-\lambda^s} [\kappa^s]^{Z_{m^S}^S} \right) \sum_{I \in \mathcal{F}_n(\mathbb{L})} \sum_{\theta^S \in \Theta^S} 1_I(\ell) \omega_{I,+} \eta_{I,Z_{m^S}^S}(\ell; \theta^S) q_I(\mathbf{x}, Z_{m^S}^S; \theta^S) \tag{27}$$

where $q_I(\mathbf{x}, Z_{m^S}^S; \theta^S)$ is the marginal density of $q(\mathbf{X}, Z_{m^S}^S)$ over any of $\mathbf{x}_{1:n}$ given the label set I and association mapping θ^S ,

$$q_I(\mathbf{x}, Z_{m^S}^S; \theta^S) = \sum_{Y \subseteq S} p_d^{Y,S}(\mathbf{x}) p_{I,+}(\mathbf{x}) \prod_{s \in Y} g^s(z_{\theta^s(\ell)}^s | \mathbf{x}) \tag{28}$$

$$\eta_{I,Z_{m^S}^S}(\ell, \theta^S) = \left\langle p_{I,+}(\cdot, \ell'), \prod_{s \in S} \psi_{Z_{m^S}^S}^s(\cdot, \ell'; \theta^s - \theta^s(\ell)) \right\rangle^{I-\ell} \tag{29}$$

where $\theta^s - \theta^s(\ell)$ denotes the remaining association mapping in θ^S except for $\theta^s(\ell)$.

MAP detection criterion determines $|\hat{\mathbf{X}}(Z_{m^S}^S)| = \hat{n}$ ($\hat{n} = 0, 1, \dots, \infty$) if and only if $Z_{m^S}^S \subseteq \mathbb{Z}_{\hat{n},m^S}^S$,

$$\mathbb{Z}_{\hat{n},m^S}^S = \left\{ Z_{m^S}^S \subseteq \mathbb{Z}_{m^S}^S : \hat{n} = \underset{n}{\operatorname{argmax}} \left(P(|\mathbf{X}| = n | Z_{m^S}^S) \right) \right\} \tag{30}$$

where $\mathbb{Z}_{\hat{n},m^S}^S = \mathbb{Z}_{\hat{n},m^{S_1}}^{S_1} \times \dots \times \mathbb{Z}_{\hat{n},m^{S_{|S|}}}^{S_{|S|}}$ is the joint measurement subspace of the SN set S where the target number is estimated as \hat{n} ; $\mathbb{Z}_{0,m^S}^S, \mathbb{Z}_{1,m^S}^S, \dots, \mathbb{Z}_{\infty,m^S}^S$ is a partition of $\mathbb{Z}_{m^S}^S$; $P(|\mathbf{X}|=n|Z_{m^S}^S)$ is the posterior probability of $|\mathbf{X}|=n$ given $Z_{m^S}^S$. From (6), $P(|\mathbf{X}|=n|Z_{m^S}^S)$ is written as

$$P(|\mathbf{X}|=n|Z_{m^S}^S) = \sum_{I \in \mathcal{F}_n(\mathbb{L})} \omega_I(Z_{m^S}^S) \tag{31}$$

where $\omega_I(Z_{m^S}^S)$ is the existing probability of the label set I given $Z_{m^S}^S$. According to the update step of Mδ-GLMB [22], $\omega_I(Z_{m^S}^S)$ is obtained as

$$\omega_I(Z_{m^S}^S) = \sum_{\theta^S \in \Theta^S} \frac{\omega_{I,+} \beta_{I,Z_{m^S}^S}(\theta^S)}{\sum_{I' \in \mathcal{F}(\mathbb{L})} \sum_{\theta'^S \in \Theta^S} \omega_{I',+} \beta_{I',Z_{m^S}^S}(\theta'^S)} \tag{32}$$

$$\beta_{I,Z_{m^S}^S}(\theta^S) = \left\langle p_{I,+}(\cdot, \ell), \prod_{s \in S} \psi_{Z_{m^S}^S}^s(\cdot, \ell; \theta^s) \right\rangle^I \tag{33}$$

where we have $\beta_{I,Z_{m^S}^S}(\theta^S) = \sum_{\ell \in \mathbb{L}_1} \eta_{I,Z_{m^S}^S}(\ell, \theta^S)$ from (29) and (33).

Let $\Psi_{\hat{n},n,m^S}$ be the integral of $q(\mathbf{X}_n, Z_{m^S}^S)$ over the space $(\mathbb{X} \times \mathbb{L})_n \times \mathbb{Z}_{\hat{n},m^S}^S$. From (18), $\Psi_{\hat{n},n,m^S}$ can be written as

$$\begin{aligned} \Psi_{\hat{n},n,m^S} &= \sum_{\ell_{1:n} \in \mathbb{L}_n} \int_{\mathbb{Z}_{\hat{n},m^S}^S} \int_{\mathbb{X}_n} q(\mathbf{X}_n, Z_{m^S}^S) dx_{1:n} dz_{1:m^S}^S \\ &= \frac{1}{\Omega_{n,m^S}} \sum_{\ell_{1:n} \in \mathbb{L}_n} \int_{\mathbb{Z}_{\hat{n},m^S}^S} \int_{\mathbb{X}_n} \prod_{s \in S} g^s(Z_{m^S}^S | \mathbf{X}_n) \pi_+(\mathbf{X}_n) dx_{1:n} dz_{1:m^S}^S \end{aligned} \tag{34}$$

(34) shows that $\Omega_{n,m^S} \Psi_{\hat{n},n,m^S} / (m^S! \cdot n!)$ is actually the probability $P(|\hat{\mathbf{X}}(Z_{m^S}^S)| = \hat{n}, |\mathbf{X}| = n, |Z^S| = m^S)$. Substituting A1 and (12) into (34) as well as using Lemma 12 in [17], $\Psi_{\hat{n},n,m^S}$ is obtained as

$$\Psi_{\hat{n},n,m^S} = \frac{n!}{\Omega_{n,m^S}} \left(\prod_{s \in S} e^{-\lambda^s} [\lambda_{\hat{n}}^s]^{m^s} \right) \sum_{I \in \mathcal{F}_n(\mathbb{L})} \sum_{\theta^S \in \Theta^S} \omega_{I,+} \phi_{I,\hat{n}}(\theta^S) \tag{35}$$

where

$$\lambda_{\hat{n}}^s = \int_{\mathbb{Z}_{\hat{n},1}^s} \kappa^s(z^s) dz^s \tag{36}$$

$$\phi_{I,\hat{n}}(\theta^S) = \left\langle p_{I,+}(\cdot, \ell), \prod_{s \in S} \left(\delta_0(\theta^s(\ell))(1 - p_d^s(\cdot, \ell)) + (1 - \delta_0(\theta^s(\ell))) \frac{p_d^s(\cdot, \ell) \int_{\mathbb{Z}_{\hat{n},1}^s} g^s(z^s | \cdot, \ell) dz^s}{\lambda_{\hat{n}}^s} \right) \right\rangle^I \tag{37}$$

Similar with $\phi_I(\theta^S)$, $\phi_{I,\hat{n}}(\theta^S)$ can be rewritten as

$$\phi_{I,\hat{n}}(\theta^S) = \left(\sum_{Y \subseteq S} \frac{\left\langle p_{I,+}(\cdot, \ell) \prod_{s \in Y} \int_{\mathbb{Z}_{\hat{n},1}^s} g^s(z^s | \cdot, \ell) dz^s, p_d^{Y,S}(\cdot, \ell) \right\rangle}{\prod_{s \in Y} \lambda_{\hat{n}}^s} \right)^I = \phi_{I,\hat{n}}(S) \tag{38}$$

(38) shows that $\phi_{I,\hat{n}}(\theta^S)$ is independent of the association mapping θ^S . From (14), (35) and (38), $\Psi_{\hat{n},n,m^S}$ is finally obtained as

$$\Psi_{\hat{n},n,m^S} = \frac{n!}{\Omega_{n,m^S}} \left(\prod_{s \in S} e^{-\lambda^s} [\lambda_{\hat{n}}^s]^{m^s} \chi_{n,m^S} \right) \sum_{I \in \mathcal{F}_n(\mathbb{L})} \omega_{I,+} \phi_{I,\hat{n}}(S) \tag{39}$$

Theorem 1: Given **A1**, **A2** and the SN set S , the lower bound for the 2nd-order LA metric based MSE of (16) is

$$\sigma_S^2 = \sum_{m^S=0}^{\infty} \sum_{n=0}^{\infty} \sum_{\hat{n}=0, n+\hat{n}>0}^{\infty} \sum_{k=0}^{\min(n,\hat{n})} \sum_{\ell \in \mathbb{L}_1} \frac{\Omega_{n,m^S} \Psi_{\hat{n},n,m^S}}{m^S!(n-k)!} \left(\varepsilon_{k,\hat{n},n} \min \left(c^2, \frac{1}{\Psi_{\hat{n},n,m^S}} \sum_{l=1}^L [J_{\hat{n},n,m^S}^{-1}(\ell)]^{l,l} \right) + (1 - \varepsilon_{k,\hat{n},n}) c^2 \right) \tag{40}$$

where c is the cut-off of the LA metric, L is the dimension of x , Ω_{n,m^S} and $\Psi_{\hat{n},n,m^S}$ are given in (24) and (39),

$$\varepsilon_{k,\hat{n},n} = \frac{k}{n + \hat{n} - k}, \quad k = 0, 1, \dots, \min(n, \hat{n}) \tag{41}$$

is actually the possible ratio of the number of common labels to the number of all distinct labels for the multi-target states and their estimates given $|\hat{\mathbf{X}}(Z_{m^S}^S)| = \hat{n}, |\mathbf{X}| = n, |Z^S| = m^S$ (that is, $|\mathcal{L}(\mathbf{X}_n) \cap \mathcal{L}(\hat{\mathbf{X}}_{\hat{n}}(Z_{m^S}^S))| / |\mathcal{L}(\mathbf{X}_n) \cup \mathcal{L}(\hat{\mathbf{X}}_{\hat{n}}(Z_{m^S}^S))|$), $J_{\hat{n},n,m^S}(\ell)$ is the $L \times L$ FIM for a single-target state with the label ℓ given $|\hat{\mathbf{X}}(Z_{m^S}^S)| = \hat{n}, |\mathbf{X}| = n, |Z^S| = m^S$,

$$[J_{\hat{n},n,m^S}(\ell)]^{i,j} = -\frac{1}{\Psi_{\hat{n},n,m^S}^2} \int_{\mathbb{Z}_{\hat{n},m^S}^S} \int_{\mathbb{X}_1} \Phi_n(x, \ell, Z_{m^S}^S) dx dz_{1:m^S}^S \quad i, j = 1, \dots, L \tag{42}$$

where the integral region $\mathbb{Z}_{\hat{n},m^S}^S$ for measurement is given by (30), $J_{\hat{n},n,m^S}(\ell) = \infty$ if $\mathbb{Z}_{\hat{n},m^S}^S = \emptyset$, the integrand $\Phi_n(x, \ell, Z_{m^S}^S)$ is

$$\begin{aligned} \Phi_n(x, \ell, Z_{m^S}^S) &= q_n(x, \ell, Z_{m^S}^S) \frac{\partial^2 \log q_n(x, \ell, Z_{m^S}^S)}{\partial x^i \partial x^j} \\ &= -\frac{1}{q_n(x, \ell, Z_{m^S}^S)} \cdot \frac{\partial q_n(x, \ell, Z_{m^S}^S)}{\partial x^i} \cdot \frac{\partial q_n(x, \ell, Z_{m^S}^S)}{\partial x^j} + \frac{\partial^2 q_n(x, \ell, Z_{m^S}^S)}{\partial x^i \partial x^j} \end{aligned} \tag{43}$$

where $q_n(x, \ell, Z_{m^S}^S)$ is given in (27).

See Appendix B for proof of Theorem 1.

Remark 2. The number of estimated targets is assumed to be unknown in the derivation of the proposed bound. Only the MAP rule, rather than the specific (or exact) number of estimated targets, is required to obtain the bound. The symbol \hat{n} used for calculating the bound in Theorem 1 is just an index for all possible (or unknown) number of estimated targets. This is similar with the symbols ℓ, k, n and m^S in Theorem 1, which are just the indices for the labels, the number of common labels in true targets and their estimates, the number of true targets and the number of sensor measurements, respectively.

Furthermore, the reason for imposing the MAP rule has been explained in **A2**. (30) has also shown that the measurement space $\mathbb{Z}_{m^S}^S$ can be divided into $\mathbb{Z}_{0,m^S}^S, \mathbb{Z}_{1,m^S}^S, \dots, \mathbb{Z}_{\infty,m^S}^S$ according to the MAP rule. It is very helpful for the proof of Theorem 1.

Remark 3. In general, the maximum number of targets and measurements can be presented by prior knowledge. Moreover, the label space $\mathbb{L}_1 = \mathbb{L}'_1 \cup \mathbb{B}_1$, where \mathbb{L}'_1 and \mathbb{B}_1 are the label spaces for the last time and new-born targets. In general, \mathbb{B}_1 can also be preseted by prior knowledge. According to these presets, the sum of infinite terms in (40) becomes the sum of finite terms.

Remark 4. Once the specific forms of $p_{I,+}(\mathbf{x})$, $p_d(\mathbf{x})$ and $g^s(z^s|\mathbf{x})$ are given, $\partial q_n(\mathbf{x}, \ell, Z_{m^s}^S)/\partial \mathbf{x}^i$ and $\partial^2 q_n(\mathbf{x}, \ell, Z_{m^s}^S)/\partial \mathbf{x}^i \partial \mathbf{x}^j$ in $\Phi_n(\mathbf{x}, \ell, Z_{m^s}^S)$ can be obtained from (27) and (28).

Remark 5. The formulas of $\Psi_{\hat{n},n,m^s}$ and $J_{\hat{n},n,m^s}(\ell)$ contain the integral over the measurement subspace $\mathbb{Z}_{\hat{n},m^s}^S$, which is calculated via MC integration [45]. To improve computational efficiency, the samples of MC integration are selected as predicted ideal measurement sets (PIMS) [31]. The calculation steps are shown in Algorithm 2.

Algorithm 2. Steps for calculating $\Psi_{\hat{n},n,m^s}$ and $J_{\hat{n},n,m^s}(\ell)$.

- 1. Prediction sampling:** Generate M samples $\tilde{\mathbf{X}}_{n,+}^{(1)}, \dots, \tilde{\mathbf{X}}_{n,+}^{(M)}$ of multi-target state sets from the predicted density $\pi_+(\mathbf{X})$;
 - 2. PIMS generating:** For $j = 1, \dots, M$, generate PIMS $\tilde{Z}_{m^s}^{S,(j)}$ of the SN set S based on $\tilde{\mathbf{X}}_{n,+}^{(j)}$ [31];
 - 3. PIMS partitioning:** Divide PIMS $\{\tilde{Z}_{m^s}^{S,(j)}\}_{j=1}^M$ into the measurement subspace $\mathbb{Z}_{\hat{n},m^s}^S$ ($\hat{n} = 0, 1, \dots, \infty$) according to (30);
 - 4. MC integration:** Given the PIMS assigned to $\mathbb{Z}_{\hat{n},m^s}^S$, $\Psi_{\hat{n},n,m^s}$ and $J_{\hat{n},n,m^s}(\ell)$ are obtained by applying MC integral formula [45] to (39) and (42).
-

4.2. SMC and GM Implementations for the Bound

In order to derive the SMC implementation of the bound in Theorem 1, it is assumed that

A3: Each $p_{I,+}(\mathbf{x})$ involved in the predicted M δ -GLMB density $\pi_+ = \{(\omega_{I,+}, p_{I,+})\}_{I \in \mathcal{F}(\mathbb{L})}$ is described by a set of weighted particles $\left\{ \left(\tilde{v}_{I,+}^i(\ell), \tilde{\mathbf{x}}_{I,+}^i(\ell) \right) \right\}_{i=1}^{\tilde{G}_{I,+}(\ell)}$,

$$p_{I,+}(\mathbf{x}) = \sum_{i=1}^{\tilde{G}_{I,+}(\ell)} \tilde{v}_{I,+}^i(\ell) \delta_{\tilde{\mathbf{x}}_{I,+}^i(\ell)}(\mathbf{x}) \tag{44}$$

Substituting (44) into (22), (28), (29), (33) and (38), $\varphi_I(S)$, $q_I(\mathbf{x}, Z_{m^s}^S; \theta^S)$, $\eta_{I,Z_{m^s}^S}(\ell, \theta^S)$, $\beta_{I,Z_{m^s}^S}(\theta^S)$ and $\phi_{I,\hat{n}}(S)$ are rewritten as

$$\varphi_I(S) = \left(\sum_{Y \subseteq S} \frac{\sum_{i=1}^{\tilde{G}_{I,+}(\ell)} \tilde{v}_{I,+}^i(\ell) p_d^{Y,S}(\tilde{\mathbf{x}}_{I,+}^i(\ell), \ell)}{\prod_{s \in Y} \lambda^s} \right)^I \tag{45}$$

$$q_I(\mathbf{x}, Z_{m^s}^S; \theta^S) = \sum_{Y \subseteq S} \sum_{i=1}^{\tilde{G}_{I,+}(\ell)} \tilde{v}_{I,+}^i(\ell) p_d^{Y,S}(\tilde{\mathbf{x}}_{I,+}^i(\ell), \ell) \prod_{s \in Y} g^s(z_{\theta^s(\ell)}^s | \tilde{\mathbf{x}}_{I,+}^i(\ell), \ell) \tag{46}$$

$$\eta_{I,Z_{m^s}^S}(\ell, \theta^S) = \left(\sum_{i=1}^{\tilde{G}_{I,+}(\ell)} \tilde{v}_{I,+}^i(\ell) \prod_{s \in S} \psi_{Z_{m^s}^S}^s(\tilde{\mathbf{x}}_{I,+}^i(\ell), \ell'; \theta^s - \theta^s(\ell)) \right)^{I-\ell} \tag{47}$$

$$\beta_{I,Z_{m^s}^S}(\theta^S) = \left(\sum_{i=1}^{\tilde{G}_{I,+}(\ell)} \tilde{v}_{I,+}^i(\ell) \prod_{s \in S} \psi_{Z_{m^s}^S}^s(\tilde{\mathbf{x}}_{I,+}^i(\ell), \ell; \theta^s) \right)^I \tag{48}$$

$$\phi_{I,\hat{n}}(S) = \left(\sum_{Y \subseteq S} \frac{\sum_{i=1}^{\tilde{G}_{I,+}(\ell)} \tilde{v}_{I,+}^i(\ell) p_d^{Y,S}(\tilde{\mathbf{x}}_{I,+}^i(\ell), \ell) \prod_{s \in Y} \int_{\mathbb{Z}_{\hat{n},1}^S} g^s(z^s | \tilde{\mathbf{x}}_{I,+}^i(\ell), \ell) dz^s}{\prod_{s \in Y} \lambda_{\hat{n}}^s} \right)^I \tag{49}$$

Finally, the SMC forms of Ω_{n,m^S} , $q_n(\mathbf{x}, Z_{m^S}^S)$, $\Psi_{\hat{n},n,m^S}$ and $\Phi_n(x, \ell, Z_{m^S}^S)$ are respectively obtained by substituting (45)–(49) into (24), (27), (39) and (43), where $\partial q_n(x, \ell, Z_{m^S}^S) / \partial x^i$ and $\partial^2 q_n(x, \ell, Z_{m^S}^S) / \partial x^i \partial x^j$ involved in $\Phi_n(x, \ell, Z_{m^S}^S)$ are

$$\begin{cases} \frac{\partial q_n(x, \ell, Z_{m^S}^S)}{\partial x^i} = \sum_{i'=1}^{\tilde{G}_{I,+}(\ell)} \tilde{v}_{I,+}^{i'}(\ell) \frac{\partial q_n(x, \ell, Z_{m^S}^S)}{\partial x^i} \Big|_{x=\tilde{x}_{I,+}^{i'}(\ell)} \\ \frac{\partial^2 q_n(x, \ell, Z_{m^S}^S)}{\partial x^i \partial x^j} = \sum_{i'=1}^{\tilde{G}_{I,+}(\ell)} \tilde{v}_{I,+}^{i'}(\ell) \frac{\partial^2 q_n(x, \ell, Z_{m^S}^S)}{\partial x^i \partial x^j} \Big|_{x=\tilde{x}_{I,+}^{i'}(\ell)} \end{cases} \quad (50)$$

In order to derive the GM implementation of the bound in Theorem 1, it is assumed that

A4: Each $p_{I,+}(\mathbf{x})$ involved in the predicted M δ -GLMB density $\pi_+ = \{(\omega_{I,+}, p_{I,+})\}_{I \in \mathcal{F}(\mathbb{L})}$ is described by the GM form of

$$p_{I,+}(\mathbf{x}) = \sum_{i=1}^{G_{I,+}(\ell)} v_{I,+}^i(\ell) \mathcal{N}(x; \mu_{I,+}^i(\ell), \Sigma_{I,+}^i(\ell)), \text{ with } \sum_{i=1}^{G_{I,+}(\ell)} v_{I,+}^i(\ell) = 1 \quad (51)$$

where $\mathcal{N}(\cdot; \mu_{I,+}^i(\ell), \Sigma_{I,+}^i(\ell))$ denotes the Gaussian density with mean $\mu_{I,+}^i(\ell)$ and covariance matrix $\Sigma_{I,+}^i(\ell)$, $v_{I,+}^i(\ell)$ and $G_{I,+}(\ell)$ are the weights and number of GM terms.

A5: The detection probability $p_d^s(\mathbf{x})$ is independent of \mathbf{x} and the likelihood function $g^s(z^s | \mathbf{x})$ is linear Gaussian,

$$p_d^s(\mathbf{x}) = p_d^s, \quad g^s(z^s | \mathbf{x}) = \mathcal{N}(z^s; H^s x, R^s) \quad (52)$$

where H^s and R^s are the observation function and covariance matrix for measurement noise.

From **A5** and (23), we have

$$p_d^{Y,S} = \prod_{s \in Y} p_d^s \prod_{s' \in (S-Y)} (1 - p_d^{s'}) \quad (53)$$

$$\prod_{s \in Y} g^s(z^s | \mathbf{x}) = \mathcal{N}(z^Y; H^Y x, R^Y) \quad (54)$$

where

$$z^Y = \begin{bmatrix} z^{s_1} \\ \vdots \\ z^{s_{|Y|}} \end{bmatrix}; \quad H^Y = \begin{bmatrix} H^{s_1} \\ \vdots \\ H^{s_{|Y|}} \end{bmatrix}; \quad R^Y = \begin{bmatrix} R^{s_1} & & \\ & \ddots & \\ & & R^{s_{|Y|}} \end{bmatrix} \quad (55)$$

Substituting (53) into (22), $\varphi_I(S)$ is rewritten as

$$\varphi_I(S) = \left(\sum_{Y \subseteq S} \frac{p_d^{Y,S}}{\prod_{s \in Y} \lambda^s} \right)^{|I|} \quad (56)$$

Substituting (51), (53) and (54) into (28) as well as using Lemma 2 in [11], $q_I(\mathbf{x}, Z_{m^S}^S; \theta^S)$ is rewritten as

$$\begin{aligned} q_I(\mathbf{x}, Z_{m^S}^S; \theta^S) &= \sum_{Y \subseteq S} \sum_{i=1}^{G_{I,+}(\ell)} p_d^{Y,S} v_{I,+}^i(\ell) \mathcal{N}(x; \mu_{I,+}^i(\ell), \Sigma_{I,+}^i(\ell)) \mathcal{N}(z_{\theta^Y}^Y; H^Y x, R^Y) \\ &= \sum_{Y \subseteq S} \sum_{i=1}^{G_{I,+}(\ell)} p_d^{Y,S} v_{I,+}^i(\ell) \mathcal{N}(z_{\theta^Y}^Y; H^Y \mu_{I,+}^i(\ell), \Xi_{I,+}^{i,Y}(\ell)) \mathcal{N}(x; \mu_{I,Z_{m^Y}^S}^i(\ell; \theta^Y), \Sigma_i^i(\ell; Y)) \end{aligned} \quad (57)$$

where

$$\begin{cases} \Xi_{I,+}^{i,Y}(\ell) = H^Y \Sigma_{I,+}^i(\ell) [H^Y]^T + R^Y \\ \mu_{I,Z_{m^Y}^Y}^i(\ell; \theta^Y) = \mu_{I,+}^i(\ell) + \Sigma_{I,+}^i(\ell) [H^Y]^T [\Xi_{I,+}^{i,Y}(\ell)]^{-1} (z_{\theta^Y(\ell)}^Y - H^Y \mu_{I,+}^i(\ell)) \\ \Sigma_I^i(\ell; Y) = \Sigma_{I,+}^i(\ell) - \Sigma_{I,+}^i(\ell) [H^Y]^T [\Xi_{I,+}^{i,Y}(\ell)]^{-1} H^Y \Sigma_{I,+}^i(\ell) \end{cases} \quad (58)$$

Similarly, substituting (51), (53) and (54) into (29), (33) and (38), $\eta_{I,Z_{m^S}^S}(\ell; \theta^S)$, $\beta_{I,Z_{m^S}^S}(\theta^S)$ and $\phi_{I,\hat{n}}(S)$ are rewritten as

$$\eta_{I,Z_{m^S}^S}(\ell; \theta^S) = \left(\sum_{Y \subseteq S} \sum_{i=1}^{G_{I,+}(\ell)} \frac{p_d^{Y,S} v_{I,+}^i(\ell)}{\prod_{s \in Y} \kappa^s(z_{[\theta^Y - \theta^S(\ell)](\ell')})} \mathcal{N}(z_{[\theta^Y - \theta^S(\ell)](\ell')}; H^Y \mu_{I,+}^i(\ell), \Xi_{I,+}^{i,Y}(\ell)) \right)^{I-\ell} \quad (59)$$

$$\beta_{I,Z_{m^S}^S}(\theta^S) = \left(\sum_{Y \subseteq S} \sum_{i=1}^{G_{I,+}(\ell)} \frac{p_d^{Y,S} v_{I,+}^i(\ell)}{\prod_{s \in Y} \kappa^s(z_{\theta^S(\ell)})} \mathcal{N}(z_{\theta^Y(\ell)}^Y; H^Y \mu_{I,+}^i(\ell), \Xi_{I,+}^{i,Y}(\ell)) \right)^I \quad (60)$$

$$\phi_{I,\hat{n}}(S) = \left(\sum_{Y \subseteq S} \sum_{i=1}^{G_{I,+}(\ell)} \frac{p_d^{Y,S} v_{I,+}^i(\ell)}{\prod_{s \in Y} \lambda_{\hat{n}}^s} \int_{Z_{\hat{n},1}^Y} \mathcal{N}(z^Y; H^Y \mu_{I,+}^i(\ell), \Xi_{I,+}^{i,Y}(\ell)) dz^Y \right)^I \quad (61)$$

Finally, the GM forms of Ω_{n,m^S} , $q_n(\mathbf{x}, Z_{m^S}^S)$, $\Psi_{\hat{n},n,m^S}$ and $\Phi_n(x, \ell, Z_{m^S}^S)$ are respectively obtained by substituting (56)–(61) into (24), (27), (39) and (43). Obviously, they no longer contain integrals of state x and have analytic forms except for $\Psi_{\hat{n},n,m^S}$. Here $\partial q_n(x, \ell, Z_{m^S}^S) / \partial x^i$ and $\partial^2 q_n(x, \ell, Z_{m^S}^S) / \partial x^i \partial x^j$ in $\Phi_n(x, \ell, Z_{m^S}^S)$ are both linear functions of $\partial \mathcal{N}(x; \mu_{I,Z_{m^Y}^Y}^i(\ell; \theta^Y), \Sigma_I^i(\ell; Y)) / \partial x^i$ and $\partial^2 \mathcal{N}(x; \mu_{I,Z_{m^Y}^Y}^i(\ell; \theta^Y), \Sigma_I^i(\ell; Y)) / \partial x^i \partial x^j$, which can be obtained by the following formulas:

$$\begin{cases} \frac{\partial \mathcal{N}(x; \mu, \Sigma)}{\partial x^i} = -[\Sigma^{-1}(x - \mu)]^i \mathcal{N}(x; \mu, \Sigma) \\ \frac{\partial^2 \mathcal{N}(x; \mu, \Sigma)}{\partial x^i \partial x^j} = [\Sigma^{-1}(x - \mu)(x - \mu)^T [\Sigma^{-1}]^T - \Sigma^{-1}]^{ij} \mathcal{N}(x; \mu, \Sigma) \end{cases} \quad (62)$$

If the observation model is non-linear, that is,

$$g^s(z^s | \mathbf{x}) = \mathcal{N}(z^s; h^s(x), R^s) \quad (63)$$

where $h^s(x)$ is the nonlinear observation function of state x . In this case, the extended Kalman (EK) or unscented Kalman (UK) filter can be used to calculate the mean and covariance matrix of each GM term [46,47].

4.3. Sub-Optimization Based on Coordinate Descent

The computational cost of this method is composed of three parts: sensor selection, M δ -GLMB filtering and weighted KLA fusion. The latter two have been analyzed in Reference [22,23]. This paper only studies the computational cost of sensor selection and its approximate algorithm. When the number of SNs is large, it has a much more significant effect on the total amount of computation than the last two.

As shown in (15) and (17), sensor selection is actually a constrained combinatorial optimization problem. To find the optimal solution by the exhaustive search method, the objective function needs to be repeatedly calculated for $C_{|N|}^{|S|} = |N|! / (|S|!(|N|-|S|)!)$ times. Obviously, its computational cost increases with the SN number $|N|$ in the form of combination explosion. In order to reduce the computational cost, some heuristic optimization algorithms, such as genetic algorithm [39] and so

forth, is used to tackle this problem. However, the convergence speed of the heuristic algorithms will become rather slow when the objective function is relatively complex. As a result, to further improve the computational speed, coordinate descent method [37] is proposed to find a sub-optimal solution of (17). Its computational cost increases with the SN number $|N|$ in an approximate polynomial form.

Set a binary switch variable $\zeta^s \in \{0, 1\}$, $s = 1, \dots, |N|$, for each SN. $\zeta^s = 1$ indicates $s \in S$ while $\zeta^s = 0$ indicates $s \notin S$. The vector $\zeta = [\zeta^1, \dots, \zeta^{|N|}]$ is composed of the switch variables of all SNs belonging to the same LFC. Clearly, the set S is completely determined by ζ . Then, (17) is relaxed to an unconstrained optimization problem by the augmented objective function

$$F(\zeta, r) = \sigma_\zeta^2 + \omega \sum_{i=1}^l \gamma_i^{-1}(\zeta) + \frac{1}{\sqrt{\omega}} \sum_{j=1}^m v_j^2(\zeta) \tag{64}$$

where $\omega > 0$ is a barrier factor, $\sum_{i=1}^l \gamma_i^{-1}(\zeta)$ and $\sum_{j=1}^m v_j^2(\zeta)$ are inequality and equality penalty terms.

Algorithm 3 presents the iteration steps for handling the relaxed problem by using the coordinate descent method.

Algorithm 3. Coordinate descent method.

Step 1: Set initial iteration number $i = 0$, initial SN switch vector $\zeta_{(0)}$, initial barrier factor ω and its reduction coefficient $0 < C < 1$;

Step 2: From $s = 1$ to $s = |N|$, calculate $\zeta_{(i+1)}^s = \operatorname{argmin}_{\zeta^s \in \{0,1\}} F(\zeta_{(i+1)}^1, \dots, \zeta_{(i+1)}^{s-1}, \zeta^s, \zeta_{(i)}^{s+1}, \dots, \zeta_{(i)}^{|N|})$, where

$\zeta_{(i+1)}^1, \dots, \zeta_{(i+1)}^{s-1}, \zeta_{(i)}^{s+1}, \dots, \zeta_{(i)}^{|N|}$ are treated as constants;

Step 3: If $\zeta_{(i+1)} = \zeta_{(i)}$, then go to Step 4; Otherwise, set $i = i + 1$, go to Step 2;

Step 4: If $\zeta_{(i+1)} = \zeta_{(0)}$, then output $\zeta_{(i+1)}$ as the solution of (17); Otherwise, set $\omega = C\omega$, $\zeta_{(0)} = \zeta_{(i+1)}$, $i = 0$ and then go to Step 2.

In order to improve the probability to converge to the global optimum and accelerate the convergence speed for the coordinate descent method, the initial barrier factor ω and its reduction coefficient C can be appropriately selected by the methods in Reference [48], the initial SN switch vector is set as $\zeta_{(0)} = \zeta'^*$, where ζ'^* is the outputted switch vector at the last time.

4.4. Weighted KLA Fusion

A1 indicates that the posterior density of each LFC is a M δ -GLMB form of $\pi^i(\mathbf{X}|Z^{S^i}) = \left\{ \left(\omega_L^i(Z^{S^i}), p_L^i(x, \ell|Z^{S^i}) \right) \right\}_{L \in \mathcal{F}(\mathbb{L})}$, $i \in C$. Then, for the LFC $j \in C$, given the LFC subset C^j connected to it and the normalized nonnegative fusion weights $\omega^{j,i}$ ($i \in C^j$), its fused density obtained by the weighted KLA rule is still a M δ -GLMB form of $\bar{\pi}^j(\mathbf{X}) = \left\{ \left(\bar{\omega}_L^j, \bar{p}_L^j(x, \ell) \right) \right\}_{L \in \mathcal{F}(\mathbb{L})}$ [23],

$$\bar{\omega}_L^j = \frac{\prod_{i \in C^j} \left(\omega_L^i(Z^{S^i}) \right)^{\omega^{j,i}} \left[\int \prod_{i \in C^j} \left(p_L^i(x, \cdot | Z^{S^i}) \right)^{\omega^{j,i}} dx \right]^L}{\sum_{L' \subseteq \mathbb{L}} \prod_{i' \in C^j} \left(\omega_{L'}^{i'}(Z^{S^{i'}}) \right)^{\omega^{j,i'}} \left[\int \prod_{i' \in C^j} \left(p_{L'}^{i'}(x, \cdot | Z^{S^{i'}}) \right)^{\omega^{j,i'}} dx \right]^{L'}} \tag{65}$$

$$\bar{p}_L^j(x, \ell) = \frac{\prod_{i \in C^j} \left(p_L^i(x, \ell | Z^{S^i}) \right)^{\omega^{j,i}}}{\int \prod_{i' \in C^j} \left(p_{L'}^{i'}(x, \ell | Z^{S^{i'}}) \right)^{\omega^{j,i'}} dx} \tag{66}$$

$[10, 10]^T$, $[10, 25]^T$ in turn. Each LFC has $|N^j| = 50$ subordinate SNs ($j = 1, \dots, 8$). So there are in total $|N| = \dots$ SNs in the entire network. Each LFC can communicate with other LFCs within 25 km apart away from it. Finally, the directed connection set A for the LFCs and the locations for all SNs are shown in Figure 2.

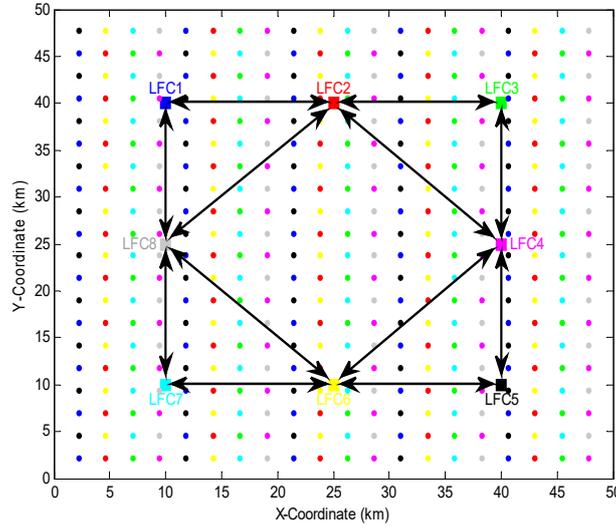


Figure 2. Locations of LFCs and SNs in the decentralized sensor network. ■ denotes LFC, ● denotes SN, different colors correspond to different LFCs and their subordinate SNs. Solid line with directed arrow indicates a communication connection between two LFCs. Each SN is 2.4 km apart away from other one.

The clutter is a uniformly distributed Poisson RFS over the region \mathcal{A} . In this example, the clutter rate and detection probability of each SN are firstly set as $\lambda^s = \lambda = 20$ and $p_d^s(x) = p_d = 0.95$, $s = 1, \dots, 400$.

To embody the performance variation of sensor observation, the network is assumed to consist of different types of SNs with distinctive observation function and noise covariance. The single-target likelihoods of the SNs are all Gaussian distributed as shown in (63). Each SN of LFC1 and LFC5 receives the distance and angle measurements of target, so its observation function $h^s(x)$ is

$$h^s(x) = \left[\|x - u^s\|, \arctan \frac{p_y - u_y^s}{p_x - u_x^s} \right]^T, \quad s \in N^1 \text{ or } s \in N^5 \quad (71)$$

where $u^s = [u_x^s, u_y^s]^T$ is the known position of the SN s , $\|x - u^s\| = \sqrt{(p_x - u_x^s)^2 + (p_y - u_y^s)^2}$ is the distance between the SN s and target. The measurement noise covariance matrix is also modeled as a nonlinear function of state x ,

$$R^s(x) = \begin{cases} \text{diag} \left(([0.2 + 0.05\|x - u^s\|] \text{km})^2, ([0.02 + 0.001\|x - u^s\|] \text{rad})^2 \right) & s \in N^1 \\ \text{diag} \left(([0.4 + 0.04\|x - u^s\|] \text{km})^2, ([0.04 + 0.0005\|x - u^s\|] \text{rad})^2 \right) & s \in N^5 \end{cases} \quad (72)$$

Each SN of LFC2 and LFC6 only receives the distance measurement of target. Its $h^s(x)$ and $R^s(x)$ are

$$h^s(x) = \|x - u^s\|, \quad s \in N^2 \text{ or } s \in N^6 \quad (73)$$

$$R^s(x) = \begin{cases} ([0.1 + 0.02\|x - u^s\|] \text{km})^2 & s \in N^2 \\ ([0.2 + 0.01\|x - u^s\|] \text{km})^2 & s \in N^6 \end{cases} \quad (74)$$

Each SN of LFC3 and LFC7 only receives the angle measurement of target. Its $h^s(x)$ and $R^s(x)$ are

$$h^s(x) = \arctan \frac{p_y - u_y^s}{p_x - u_x^s}, \quad s \in N^3 \text{ or } s \in N^7 \quad (75)$$

$$R^s(x) = \begin{cases} ([0.01 + 0.001\|x - u^s\|]\text{rad})^2 & s \in N^3 \\ ([0.02 + 0.0005\|x - u^s\|]\text{rad})^2 & s \in N^7 \end{cases} \quad (76)$$

Each SN of LFC4 and LFC8 receives distance and Doppler measurements of target. Its $h^s(x)$ and $R^s(x)$ are

$$h^s(x) = \left[\|x - u^s\|, \frac{(p_x - u_x^s)\dot{p}_x + (p_y - u_y^s)\dot{p}_y}{\|x - u^s\|} \right]^T, \quad s \in N^4 \text{ or } s \in N^8 \quad (77)$$

$$R^s(x) = \begin{cases} \text{diag} \left(([0.2 + 0.05\|x - u^s\|]\text{km})^2, ([0.02 + 0.001\|x - u^s\|]\text{km/s})^2 \right) & s \in N^4 \\ \text{diag} \left(([0.4 + 0.04\|x - u^s\|]\text{km})^2, ([0.04 + 0.0005\|x - u^s\|]\text{km/s})^2 \right) & s \in N^8 \end{cases} \quad (78)$$

In this example, there are three constraints for the sensor selection optimization of (17), which are

C1: Due to the limitations of communication bandwidth, energy consumption, computation capacity and storage space, the LFC j can only select K^j SNs at each scan ($K^j \leq |N^j|$),

$$K^j - |S^j| = 0 \quad (79)$$

In this example, if $j = 1, 4, 5, 8$, then $K^j = 5$; otherwise $K^j = 8$.

C2: The field of view (FoV) of the SN s is modeled as a circular area with the center u^s and radius ρ^s , $\mathcal{A}^s(\rho^s) = \{ p^s \in \mathcal{A} : \|p^s - u^s\| \leq \rho^s \}$, $\mathcal{A}^s \subseteq \mathcal{A}$. The FoVs of different SNs can be overlapped. To ensure that the FoVs of the SN set S^j can totally cover the region \mathcal{A} , it is required that

$$\mathcal{A} - \bigcup_{s \in S^j} \mathcal{A}^s(\rho^s) = 0 \quad (80)$$

In this example, if $j = 1, 4, 5, 8$, then $\rho^s = 30$ km; otherwise $\rho^s = 20$ km.

C3: To avoid mutual interference between the homogeneous SNs belonging to the same LFC, the distance between any two SNs in S^j must be not smaller than the threshold D^j ,

$$\min_{s, s' \in S^j} \|u^s - u^{s'}\| - D^j \geq 0 \quad (81)$$

In this example, $D^j = 5$ km for $j = 1, \dots, 8$.

According to the objective function and calculation method used for sensor selection, our algorithm is abbreviated as *LA bound with coordinate descent*. It is firstly implemented by the SMC technique and coded by MATLAB R2018a. Each $p_I(\cdot, \ell)$ involved in the M δ -GLMB density is approximated by 500 particles on average. In this example, the maximum numbers of targets and measurements of each SN per scan are set as 25 and 200, the cut-off is set as $c = 1000$ m. The algorithm is testing on a desktop with the CPU of AMD Ryzen 7 2700X and 64 GB RAM. We conduct 500 MC simulations, each of which includes $T = 25$ scans (a total of 250 s). In these simulations, the target tracks (including the instants of birth and death), clutter and measurements originating from targets are generated independently according to the aforementioned models.

We firstly present the result of sensor selection obtained by the algorithm in one simulation. Figure 3 shows the target trajectories in the simulation, where a total of 15 targets are generated at different instants and locations. For easy description, the targets are named as T1~T15 in turn. The name and survival period of each target are marked at the start point of the target. During the surveillance period, the number of targets at the initial time is the least (3 targets). The number of

targets at the 15th~18th scans is the most (13 targets). The target T1 intersects with the target T10 at the 15th scan.

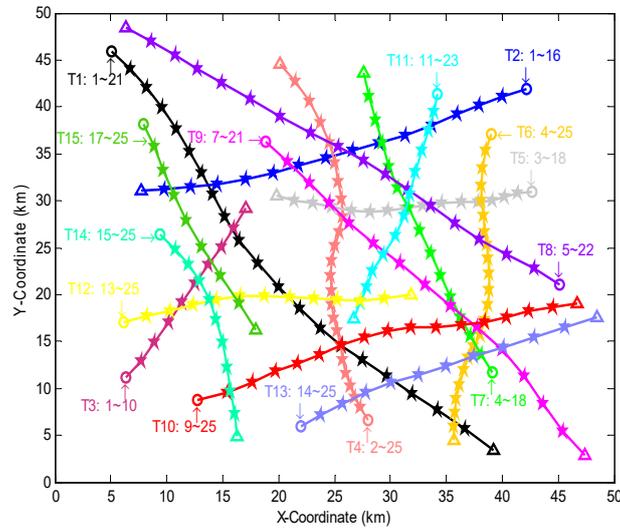
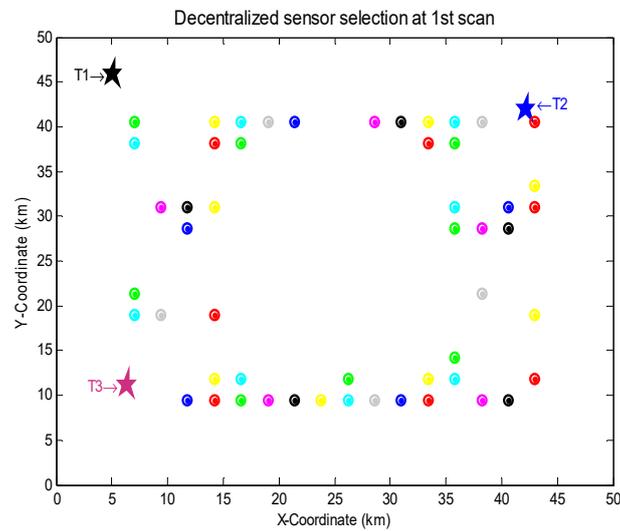


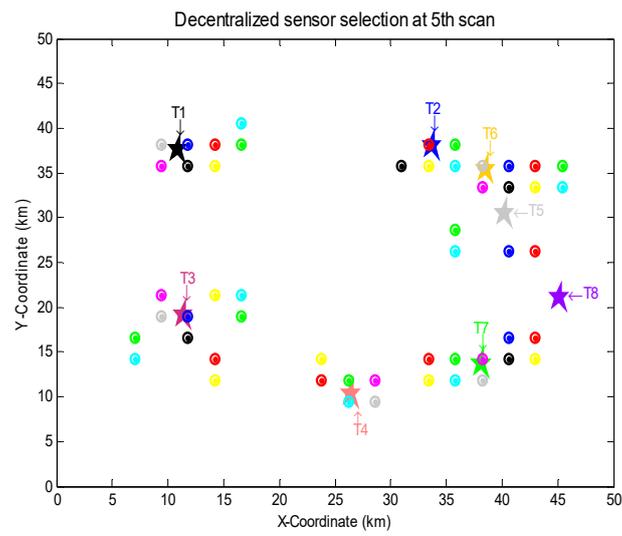
Figure 3. Target trajectories in a simulation. ○, △ and ★ are the start point, end point and rest positions of a target, the solid line is the track of a target, different colors correspond to different targets.

Figure 4a~f show the results of sensor selection at the 1st, 5th, 10th, 15th, 20th and 25th scans. It can be seen that the SNs selected by each LFC will change adaptively with the multi-target movement versus time. Specifically, in order to minimize the bound in Theorem 1, most of the selected SNs locate in the regions closer to the survival targets at each scan. A few SNs which are far from the survival targets are selected to satisfy Constraint C2. Moreover, due to Constraint C3, the homogeneous SNs of each LFC cannot be excessively concentrated in a small region.

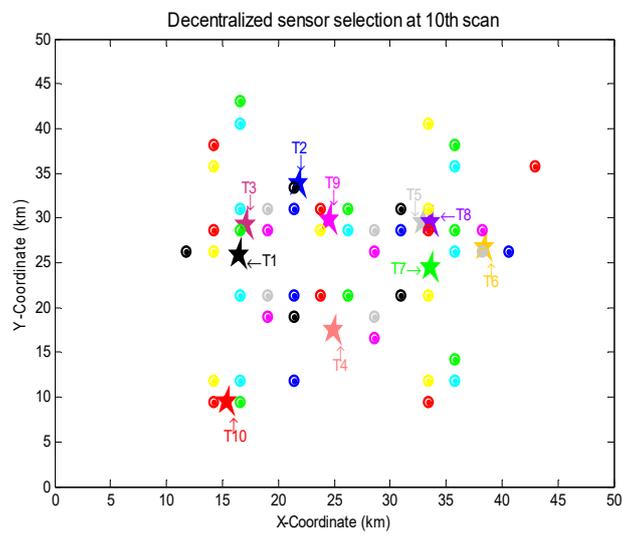


(a)

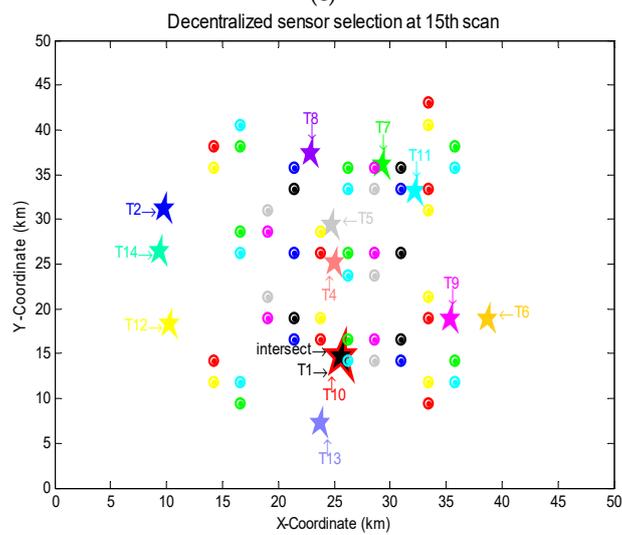
Figure 4. Cont.



(b)



(c)



(d)

Figure 4. Cont.

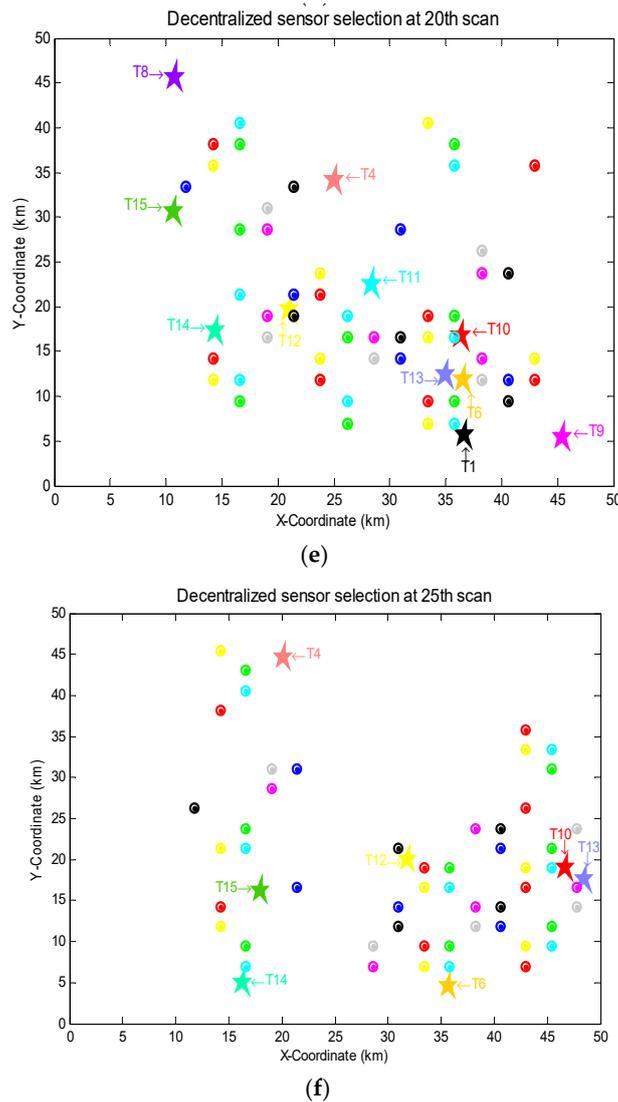


Figure 4. Sensor selection using *LA bound with coordinate descend* at (a) the 1st scan; (b) the 5th scan; (c) the 10th scan; (d) the 15th scan; (e) the 20th scan; (f) the 25th scan. ★ and ⊙ denote the targets and selected SNs, different colors correspond to different targets and SNs of different LFCs.

In order to further verify the performance of this method in tracking accuracy and computational time, it is compared with the methods of *LA bound with genetic algorithm*, *CS divergence with genetic algorithm* and *Random selection* under the same test platform. In the genetic algorithm, the population size is 50, crossing rate is 0.9, mutation rate is 0.001, elite rate is 0.04 and the maximum number of iterations is 500. In the CS divergence method, the objective function of the LFC j is

$$\left[S^j \right]^* = \operatorname{argmax}_{S^j \subseteq N^j} E \left[D_{CS} \left(\pi_+^j, \pi^j \left(\cdot | Z^{S^j} \right) \right) \right] \quad (82)$$

where $D_{CS}(\phi, \varphi)$ denotes the CS divergence between the densities ϕ and φ . [31] presents the specific form of the CS divergence when ϕ and φ are both GLMB densities. Since the posterior density $\pi^j(\mathbf{X} | Z^{S^j})$ is unknown before sensor selection, the expected value rather than real value of the CS divergence is applied in (82). In order to calculate the expected value, the MC integration based on PIMS also needs to be used here.

For comparison, both the OSPA and LA metrics are used to measure the error of multi-target position estimates.

Figure 5a,b present the 500 MC averages of the OSPA and LA errors for the four methods. Note that the error here is selected as the average of all LFCs because of Remark 1.

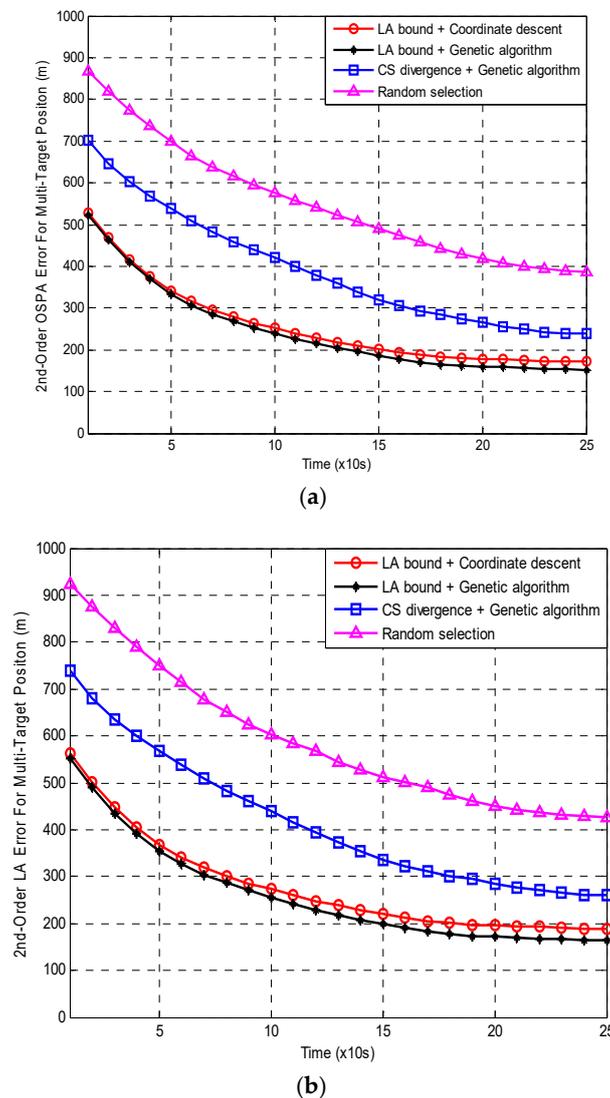


Figure 5. 500 MC averages of (a) OSPA and (b) LA errors versus time with $c = 1000$ m.

Figure 5 shows that both the averaged OSPA and LA errors from all the four methods decrease with time. Furthermore, the LA errors are always larger than the relevant OSPA errors. The reason for this has been explained in Section 2.3. In both of Figure 5a,b, the errors from *Random selection* are always the largest. The next is the errors from *CS divergence with genetic algorithm*. The errors from *LA bound with genetic algorithm* are always the smallest. The errors from *LA bound with coordinate descent* are slightly larger than those of *LA bound with genetic algorithm*. Compared with *Random selection*, the errors of the other three algorithms are approximately reduced by 40%, 60% and 55%, respectively. Obviously, the two LA bound based methods outperform the CS divergence based method in MTT accuracy. There are three reasons for this:

1) The LA bound has a clearer physical meaning than the CS divergence. This is because that the former indicates the achievable optimal MTT accuracy with labeled RFS state. In contrast, the latter is not directly related to the MTT accuracy since maximizing the CS divergence in (82) cannot guarantee to minimize the OSPA or LA error.

2) The CS divergence cannot provide a basis for setting the weights of KLA rule as the LA bound. Therefore, in the CS divergence based method, the KLA weights can only be set to the same by

convention [32]. However, in this example the MTT accuracy of different LFCs is probably not the same because of the distinguishing observation performance of the diverse SNs. If the KLA weights are set to the same without discrimination, the fusion efficiency will decline dramatically in this case.

3) In the step of sensor selection optimization, the coordinate descent method may trap in one of local optimums. In contrast, the genetic algorithm may jump out of local optimums with certain probability by its randomness. This leads that the MTT accuracy of the former is a litter worse than the latter.

On the other hand, the computational cost of a method is in general measured by its CPU run time. Then, the averaged CPU run time per scan for *LA bound with coordinate descent*, *LA bound with genetic algorithm* and *CS divergence with genetic algorithm* are 0.62s, 6.18s and 5.69s, respectively. Note that the run time here is also selected as the average of all LFCs because of Remark 1. *Random selection* obviously does not need the optimization for sensor selection, so it has no time consumption of this step. It can be seen from this that although the coordinate descent method is slightly worse than the genetic algorithm in tracking accuracy, it significantly shortens the calculation time of sensor selection optimization. In addition, the time consumption of *CS divergence with genetic algorithm* is slightly less than that of *LA bound with genetic algorithm*. This case indicates that the computational cost of the proposed bound is a little larger than that of the CS divergence expectation in (82).

In order to show the influence of different SNR on our method, the clutter rate and detection probability of each SN are changed into $(\lambda = 40, p_d = 0.85)$, $(\lambda = 60, p_d = 0.75)$, $(\lambda = 80, p_d = 0.65)$ and $(\lambda = 100, p_d = 0.55)$. Tables 1 and 2 present the final values of the OSPA and LA errors for the four methods in each scenario after 500 MC run average.

Table 1. Final value of OSPA error (Unit: m).

Clutter rate and Detection Probability	Sensor Selection Method				
	$\lambda = 20$ $p_d = 0.95$	$\lambda = 40$ $p_d = 0.85$	$\lambda = 60$ $p_d = 0.75$	$\lambda = 80$ $p_d = 0.65$	$\lambda = 100$ $p_d = 0.55$
<i>LA bound with coordinate descent</i>	171.3	193.0	218.9	249.1	283.7
<i>LA bound with genetic algorithm</i>	151.9	170.8	193.6	220.5	251.6
<i>CS divergence with genetic algorithm</i>	238.6	278.5	333.7	405.3	494.4
<i>Random selection</i>	386.7	436.0	490.5	550.1	615.2

Table 2. Final value of LA error (Unit: m).

Clutter rate and Detection Probability	Sensor Selection Method				
	$\lambda = 20$ $p_d = 0.95$	$\lambda = 40$ $p_d = 0.85$	$\lambda = 60$ $p_d = 0.75$	$\lambda = 80$ $p_d = 0.65$	$\lambda = 100$ $p_d = 0.55$
<i>LA bound with coordinate descent</i>	188.8	223.9	255.0	288.6	323.9
<i>LA bound with genetic algorithm</i>	165.7	199.6	229.5	262.9	292.3
<i>CS divergence with genetic algorithm</i>	260.1	321.2	383.9	463.9	562.8
<i>Random selection</i>	426.7	500.5	564.3	634.2	710.9

Moreover, the CPU times consumed by the sensor selection optimization for the first three methods are nearly the same for all the SNR scenarios. This is because that the sensor selection is irrelevant to the specific measurement realizations since it must be completed before the measurements are received.

It can be seen from Algorithm 1 and 2 that as clutter density increases and detection probability decreases,

1) The OSPA and LA errors of all the four methods increase in different degrees but the size order of them is always the same as that of Scenario 1;

2) Taking the error of *Random selection* as the benchmark, the improvement ratio of *CS divergence with genetic algorithm* is gradually reduced from about 40% to about 20%. Meanwhile, the improvement ratios of the two LA bound based methods are, respectively, maintained at about 60% and 55%.

The two points reflect that the lower the SNR is, the worse the sensor selection efficiency and tracking accuracy of the CS divergence based methods become. By contrast, the tracking accuracy of the LA bound based methods always maintains a good improvement ratio for all the SNR scenarios.

Assuming that the simulation scenarios remain unchanged, the above four methods are re-implemented by the GM technique where the non-linear likelihood of each SN is approximated by the EK filter,

$$\hat{H}^s \approx \left. \frac{\partial h^s(x)}{\partial x} \right|_{x=\hat{x}_+}; \hat{R}^s \approx R^s(\hat{x}_+) \quad (83)$$

The pruning and merging technology [11] are used to manage the GM terms. Let the number of the GM terms approximating to each $p_I(\cdot, \ell)$ be no more than 30. The thresholds for merging, pruning and state extraction are set to 4, 10^{-4} and 0.5, respectively. The simulation results of the GM implementation are basically consistent with those of the SMC implementation, except that the OSPA or LA error increases by about 8%. But the calculation time for sensor selection decreases by about 70%.

6. Conclusions and Future Work

A sensor selection optimization algorithm is proposed for the decentralized large-scale MTT network under the labeled RFS framework. The LA metric defined in this paper is used to measure the error between the labeled multi-target states and their estimates. The lower bound of the LA metric based MSE is taken as the cost function of sensor selection. The bound is derived by the information inequality and then, implemented by the SMC or GM technique. Then, the coordinate descent method is used to reduce the computational cost of sensor selection. Simulation results show that when the sensors of the decentralized network have different observation performance, our method outperforms the CS divergence based sensor selection algorithm in MTT accuracy.

Our future work will focus on the following two aspects:

1) Extend the proposed method to the cases of asynchronous measurement or correlated measurement noise;

2) Reference [50] has presented a very efficient implementation of the GLMB filter with linear complexity in the number of measurements and this filter has been demonstrated to handle over 1 million tracks simultaneously [51]. Therefore, it would be very helpful to improve our current study by the use of the methods proposed in Reference [50,51].

Author Contributions: Conceptualization, F.L.; Methodology, F.L.; Software, L.H.; Validation, F.L., L.H., B.W. and C.H.; Formal Analysis, B.W.; Investigation, C.H.; Resources, B.W.; Data Curation, F.L.; Writing-Original Draft Preparation, F.L.; Writing-Review & Editing, L.H.; Visualization, B.W.; Supervision, C.H.; Project Administration, F.L.; Funding Acquisition, F.L.

Funding: This research was supported by the National Natural Science Foundation of China (61473217) and the National Key Fundamental Research & Development Programs (973) of China (2013CB329405).

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Proof that $\bar{d}_p^{(c)}(\mathbf{X}, \mathbf{Y})$ is a metric

A mapping $d : X \times X \rightarrow [0, \infty)$ is called a metric if it obeys the following three properties

- 1) **Identity:** $d(x, y) = 0$ if and only if $x = y$;
- 2) **Symmetry:** $d(x, y) = d(y, x)$ for all $x, y \in X$;
- 3) **Triangle inequality:** $d(x, y) \leq d(x, z) + d(z, y)$ for all $x, y, z \in X$.

It can be obtained easily from the definition of $\bar{d}_p^{(c)}(\mathbf{X}, \mathbf{Y})$ in (9) that $0 \leq \bar{d}_p^{(c)}(\mathbf{X}, \mathbf{Y}) \leq c$ and $\bar{d}_p^{(c)}(\mathbf{X}, \mathbf{Y})$ obeys the properties of identity and symmetry. Next, we will prove the triangle inequality

$$\bar{d}_p^{(c)}(\mathbf{X}, \mathbf{Y}) \leq \bar{d}_p^{(c)}(\mathbf{X}, \mathbf{Z}) + \bar{d}_p^{(c)}(\mathbf{Z}, \mathbf{Y}) \quad (A1)$$

Let $\mathbf{x}, \mathbf{y}, \mathbf{z}$ and $\mathcal{L}(\mathbf{X}), \mathcal{L}(\mathbf{Y}), \mathcal{L}(\mathbf{Z})$ denote the individual elements and label sets of $\mathbf{X}, \mathbf{Y}, \mathbf{Z}$, respectively. We firstly set

$$x_\ell := u_\ell \quad \text{for } \ell \in \mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Y}) \cup \mathcal{L}(\mathbf{Z}) - \mathcal{L}(\mathbf{X}) \quad (\text{A2})$$

$$y_\ell := v_\ell \quad \text{for } \ell \in \mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Y}) \cup \mathcal{L}(\mathbf{Z}) - \mathcal{L}(\mathbf{Y}) \quad (\text{A3})$$

$$\begin{cases} z_\ell := u_\ell & \text{for } \ell \in \mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Y}) \cup \mathcal{L}(\mathbf{Z}) - \mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Z}) \\ z_\ell := v_\ell & \text{for } \ell \in \mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Z}) - \mathcal{L}(\mathbf{Z}) \end{cases} \quad (\text{A4})$$

where u_ℓ and v_ℓ satisfy

$$\|u_\ell - w\| \geq c, \|v_\ell - w\| \geq c \text{ and } \|u_\ell - v_{\ell'}\| \geq c \quad (\text{A5})$$

for all $\mathbf{w} \in \mathbf{X} \cup \mathbf{Y} \cup \mathbf{Z}$ and all $\ell, \ell' \in \mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Y}) \cup \mathcal{L}(\mathbf{Z})$.

Based on the setting of (A2)–(A5), the definition of $\bar{d}_p^{(c)}(\mathbf{X}, \mathbf{Y})$ in (9) is rewritten as

$$\bar{d}_p^{(c)}(\mathbf{X}, \mathbf{Y}) = \left(\frac{\sum_{\ell \in \mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Y})} d^{(c)}(x_\ell, y_\ell)^p}{|\mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Y})|} \right)^{\frac{1}{p}} \quad (\text{A6})$$

We will prove that the triangle inequality of (A1) holds for all six possible ranks of $|\mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Y})|$, $|\mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Z})|$ and $|\mathcal{L}(\mathbf{Y}) \cup \mathcal{L}(\mathbf{Z})|$.

Case 1 ($|\mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Y})| \leq |\mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Z})| \leq |\mathcal{L}(\mathbf{Y}) \cup \mathcal{L}(\mathbf{Z})|$): Since the cut-off distance $d^{(c)}(x_\ell, y_\ell) \leq c$ for all $\ell \in \mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Y}) \cup \mathcal{L}(\mathbf{Z})$ and $|\mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Y})| \leq |\mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Z})|$, (A6) can be amplified to

$$\begin{aligned} \bar{d}_p^{(c)}(\mathbf{X}, \mathbf{Y}) &\leq \left(\frac{\sum_{\ell \in \mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Z})} d^{(c)}(x_\ell, y_\ell)^p}{|\mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Z})|} \right)^{\frac{1}{p}} \\ &\leq \left(\frac{\sum_{\ell \in \mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Z})} (d^{(c)}(x_\ell, z_\ell) + d^{(c)}(z_\ell, y_\ell))^p}{|\mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Z})|} \right)^{\frac{1}{p}} \end{aligned} \quad (\text{A7})$$

where the last line holds because $d^{(c)}(x_\ell, y_\ell)$ obeys the triangle inequality.

Applying Minkowski's inequality to (A7), we get

$$\begin{aligned} \bar{d}_p^{(c)}(\mathbf{X}, \mathbf{Y}) &\leq \left(\frac{\sum_{\ell \in \mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Z})} d^{(c)}(x_\ell, z_\ell)^p}{|\mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Z})|} \right)^{\frac{1}{p}} + \left(\frac{\sum_{\ell \in \mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Z})} d^{(c)}(z_\ell, y_\ell)^p}{|\mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Z})|} \right)^{\frac{1}{p}} \\ &= \bar{d}_p^{(c)}(\mathbf{X}, \mathbf{Z}) + \left(\frac{\sum_{\ell \in \mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Z})} d^{(c)}(z_\ell, y_\ell)^p}{|\mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Z})|} \right)^{\frac{1}{p}} \end{aligned} \quad (\text{A8})$$

Similar with the derivation of the first line in (A7), the second term on the right-hand side of '=' in (A8) can be amplified to

$$\begin{aligned} \left(\frac{\sum_{\ell \in \mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Z})} d^{(c)}(z_\ell, y_\ell)^p}{|\mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Z})|} \right)^{\frac{1}{p}} &\leq \left(\frac{\sum_{\ell \in \mathcal{L}(\mathbf{Y}) \cup \mathcal{L}(\mathbf{Z})} d^{(c)}(z_\ell, y_\ell)^p}{|\mathcal{L}(\mathbf{Y}) \cup \mathcal{L}(\mathbf{Z})|} \right)^{\frac{1}{p}} \\ &= \bar{d}_p^{(c)}(\mathbf{Y}, \mathbf{Z}) \end{aligned} \quad (\text{A9})$$

because of $d^{(c)}(x_\ell, y_\ell) \leq c$ and $|\mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Z})| \leq |\mathcal{L}(\mathbf{Y}) \cup \mathcal{L}(\mathbf{Z})|$.

Finally, the triangle inequality of (A1) for Case 1 is obtained by substituting (A9) into (A8).

Case 2 ($|\mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Y})| \leq |\mathcal{L}(\mathbf{Y}) \cup \mathcal{L}(\mathbf{Z})| \leq |\mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Z})|$):

The proof of Case 2 is similar with that of Case 1 except that its first amplification is completed by using $|\mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Y})| \leq |\mathcal{L}(\mathbf{Y}) \cup \mathcal{L}(\mathbf{Z})|$ and its last amplification is completed by using $|\mathcal{L}(\mathbf{Y}) \cup \mathcal{L}(\mathbf{Z})| \leq |\mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Z})|$.

Case 3 ($|\mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Z})| \leq |\mathcal{L}(\mathbf{Y}) \cup \mathcal{L}(\mathbf{Z})| \leq |\mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Y})|$):

Using the triangle inequality of $d^{(c)}(x_\ell, y_\ell)$, (A6) can be amplified to

$$\begin{aligned} \bar{d}_p^{(c)}(\mathbf{X}, \mathbf{Y}) &\leq \left(\frac{\sum_{\ell \in \mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Y})} (d^{(c)}(x_\ell, z_\ell) + d^{(c)}(z_\ell, y_\ell))^p}{|\mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Y})|} \right)^{\frac{1}{p}} \\ &\leq \left(\frac{\sum_{\ell \in \mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Y})} (d^{(c)}(x_\ell, z_\ell))^p}{|\mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Y})|} \right)^{\frac{1}{p}} + \left(\frac{\sum_{\ell \in \mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Y})} (d^{(c)}(z_\ell, y_\ell))^p}{|\mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Y})|} \right)^{\frac{1}{p}} \end{aligned} \tag{A10}$$

where the last line is derived from Minkowski’s inequality.

Because of $|\mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Z})| \leq |\mathcal{L}(\mathbf{Y}) \cup \mathcal{L}(\mathbf{Z})| \leq |\mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Y})|$, (A10) can be amplified to

$$\bar{d}_p^{(c)}(\mathbf{X}, \mathbf{Y}) \leq \left(\frac{\sum_{\ell \in \mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Y})} (d^{(c)}(x_\ell, z_\ell))^p}{|\mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Z})|} \right)^{\frac{1}{p}} + \left(\frac{\sum_{\ell \in \mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Y})} (d^{(c)}(z_\ell, y_\ell))^p}{|\mathcal{L}(\mathbf{Z}) \cup \mathcal{L}(\mathbf{Y})|} \right)^{\frac{1}{p}} \tag{A11}$$

Then, based on the setting of (A2)–(A5), we have

$$\begin{cases} \sum_{\ell \in \mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Y})} (d^{(c)}(x_\ell, z_\ell))^p \leq \sum_{\ell \in \mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Z})} (d^{(c)}(x_\ell, z_\ell))^p \\ \sum_{\ell \in \mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Y})} (d^{(c)}(z_\ell, y_\ell))^p \leq \sum_{\ell \in \mathcal{L}(\mathbf{Z}) \cup \mathcal{L}(\mathbf{Y})} (d^{(c)}(z_\ell, y_\ell))^p \end{cases} \tag{A12}$$

Substituting (A12) into (A11), (A11) can be amplified to

$$\begin{aligned} \bar{d}_p^{(c)}(\mathbf{X}, \mathbf{Y}) &\leq \left(\frac{\sum_{\ell \in \mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Z})} (d^{(c)}(x_\ell, z_\ell))^p}{|\mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Z})|} \right)^{\frac{1}{p}} + \left(\frac{\sum_{\ell \in \mathcal{L}(\mathbf{Z}) \cup \mathcal{L}(\mathbf{Y})} (d^{(c)}(z_\ell, y_\ell))^p}{|\mathcal{L}(\mathbf{Z}) \cup \mathcal{L}(\mathbf{Y})|} \right)^{\frac{1}{p}} \\ &= \bar{d}_p^{(c)}(\mathbf{X}, \mathbf{Z}) + \bar{d}_p^{(c)}(\mathbf{Z}, \mathbf{Y}) \end{aligned} \tag{A13}$$

which is just the triangle inequality of (A1) for Case 3.

Case 4 ($|\mathcal{L}(\mathbf{Y}) \cup \mathcal{L}(\mathbf{Z})| \leq |\mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Z})| \leq |\mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Y})|$):

The proof of Case 4 is exactly the same as that of Case 3.

Case 5 ($|\mathcal{L}(\mathbf{Y}) \cup \mathcal{L}(\mathbf{Z})| \leq |\mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Y})| \leq |\mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Z})|$):

First, (A10) still holds for Case 5 depending on the triangle inequality of $d^{(c)}(x_\ell, y_\ell)$ and Minkowski’s inequality.

Since $d^{(c)}(x_\ell, y_\ell) \leq c$ for all $\ell \in \mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Y}) \cup \mathcal{L}(\mathbf{Z})$ and $|\mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Y})| \leq |\mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Z})|$, we have

$$\begin{aligned} \left(\frac{\sum_{\ell \in \mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Y})} (d^{(c)}(x_\ell, z_\ell))^p}{|\mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Y})|} \right)^{\frac{1}{p}} &\leq \left(\frac{\sum_{\ell \in \mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Z})} (d^{(c)}(x_\ell, z_\ell))^p}{|\mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Z})|} \right)^{\frac{1}{p}} \\ &= \bar{d}_p^{(c)}(\mathbf{X}, \mathbf{Z}) \end{aligned} \tag{A14}$$

From $|\mathcal{L}(\mathbf{Y}) \cup \mathcal{L}(\mathbf{Z})| \leq |\mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Y})|$, we have

$$\begin{aligned} \left(\frac{\sum_{\ell \in \mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Y})} (d^{(c)}(z_\ell, y_\ell))^p}{|\mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Y})|} \right)^{\frac{1}{p}} &\leq \left(\frac{\sum_{\ell \in \mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Y})} (d^{(c)}(z_\ell, y_\ell))^p}{|\mathcal{L}(\mathbf{Z}) \cup \mathcal{L}(\mathbf{Y})|} \right)^{\frac{1}{p}} \\ &\leq \left(\frac{\sum_{\ell \in \mathcal{L}(\mathbf{Z}) \cup \mathcal{L}(\mathbf{Y})} (d^{(c)}(z_\ell, y_\ell))^p}{|\mathcal{L}(\mathbf{Z}) \cup \mathcal{L}(\mathbf{Y})|} \right)^{\frac{1}{p}} \\ &= \bar{d}_p^{(c)}(\mathbf{Z}, \mathbf{Y}) \end{aligned} \tag{A15}$$

where the second line is obtained due to (A12).

Substituting (A14) and (A15) into (A10), we finally obtain the triangle inequality of (A1) for Case 5.

Case 6 ($|\mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Z})| \leq |\mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Y})| \leq |\mathcal{L}(\mathbf{Y}) \cup \mathcal{L}(\mathbf{Z})|$):

The proof of Case 6 is similar with that of Case 5, except that (A14) is derived from $|\mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Z})| \leq |\mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Y})|$ and (A12) while (A15) is derived from $d^{(c)}(x_\ell, y_\ell) \leq c$ and $|\mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Y})| \leq |\mathcal{L}(\mathbf{Y}) \cup \mathcal{L}(\mathbf{Z})|$.

Above all, the proof that $\bar{d}_p^{(c)}(\mathbf{X}, \mathbf{Y})$ in (9) is a metric has been completed. \square

Appendix B

Proof of Theorem 1

From (4) and (18), σ_S^2 in (16) is written as

$$\sigma_S^2 = \sum_{m^S=0}^{\infty} \sum_{n=0}^{\infty} \frac{\Omega_{n,m^S}}{m^S!n!} \sum_{\ell_{1:n} \in \mathbb{L}_n} \int_{\mathbb{Z}_{m^S}^S} \int_{\mathbb{X}_n} q(\mathbf{X}_n, Z_{m^S}^S) e^2(\mathbf{X}_n, \hat{\mathbf{X}}(Z_{m^S}^S)) dx_{1:n} dz_{1:m^S}^S \tag{A16}$$

where $q(\mathbf{X}_n, Z_{m^S}^S)$ and Ω_{n,m^S} are defined in (18) and (19), Ω_{n,m^S} is finally obtained as (24).

Dividing the integral region $\mathbb{Z}_{m^S}^S$ of (A16) into $\mathbb{Z}_{0,m^S}^S, \mathbb{Z}_{1,m^S}^S, \dots, \mathbb{Z}_{\infty,m^S}^S$ according to (30), σ_S^2 is rewritten as

$$\sigma_S^2 = \sum_{m^S=0}^{\infty} \sum_{n=0}^{\infty} \frac{\Omega_{n,m^S}}{m^S!n!} \sum_{\hat{n}=0}^{\infty} \sum_{\ell_{1:n} \in \mathbb{L}_n} \int_{\mathbb{Z}_{\hat{n},m^S}^S} \int_{\mathbb{X}_n} q(\mathbf{X}_n, Z_{m^S}^S) e^2(\mathbf{X}_n, \hat{\mathbf{X}}_{\hat{n}}(Z_{m^S}^S)) dx_{1:n} dz_{1:m^S}^S \tag{A17}$$

where the error $e(\mathbf{X}_n, \hat{\mathbf{X}}_{\hat{n}}(Z_{m^S}^S))$ is measured by the 2nd-order LA metric $\bar{d}_2^{(c)}(\mathbf{X}_n, \hat{\mathbf{X}}_{\hat{n}}(Z_{m^S}^S))$ defined in (9). Substituting $\bar{d}_2^{(c)}(\mathbf{X}_n, \hat{\mathbf{X}}_{\hat{n}}(Z_{m^S}^S))$ into (A17) and then using the identical equation $|\mathcal{L}(\mathbf{X}) \cup \mathcal{L}(\mathbf{Y})| = |\mathcal{L}(\mathbf{X})| + |\mathcal{L}(\mathbf{Y})| - |\mathcal{L}(\mathbf{X}) \cap \mathcal{L}(\mathbf{Y})|$, we get

$$\sigma_S^2 = \sum_{m^S=0}^{\infty} \sum_{n=0}^{\infty} \frac{\Omega_{n,m^S}}{m^S!n!} \sum_{\hat{n}=0, \hat{n}>0}^{\infty} \sum_{\ell_{1:n} \in \mathbb{L}_n} \frac{1}{n+\hat{n}-|\mathcal{L}(\mathbf{X}_n) \cap \mathcal{L}(\hat{\mathbf{X}}_{\hat{n}}(Z_{m^S}^S))|} \int_{\mathbb{Z}_{\hat{n},m^S}^S} \int_{\mathbb{X}_n} q(\mathbf{X}_n, Z_{m^S}^S) \cdot \left(\sum_{\ell' \in \mathcal{L}(\mathbf{X}_n) \cap \mathcal{L}(\hat{\mathbf{X}}_{\hat{n}}(Z_{m^S}^S))} \min(c^2, \|x_{\ell'} - \hat{x}_{\ell'}(Z_{m^S}^S)\|^2) + c^2(n+\hat{n}-2|\mathcal{L}(\mathbf{X}_n) \cap \mathcal{L}(\hat{\mathbf{X}}_{\hat{n}}(Z_{m^S}^S))|) \right) dx_{1:n} dz_{1:m^S}^S \tag{A18}$$

where $\ell_{1:n}$ are the labels of \mathbf{X}_n . It is obvious that $\mathcal{L}(\mathbf{X}_n) \cap \mathcal{L}(\hat{\mathbf{X}}_{\hat{n}}(Z_{m^S}^S)) \subseteq \{\ell_{1:n}\}$.

Let $k = |\mathcal{L}(\mathbf{X}_n) \cap \mathcal{L}(\hat{\mathbf{X}}_{\hat{n}}(Z_{m^S}^S))|$. $0 \leq k \leq \min(n, \hat{n})$. Given k and $\ell_{1:n}$, there are $C_n^k = n!/(k!(n-k)!)$ possible choices from $\ell_{1:n}$ for the elements of $\{\ell'_{1:k}\} = \mathcal{L}(\mathbf{X}_n) \cap \mathcal{L}(\hat{\mathbf{X}}_{\hat{n}}(Z_{m^S}^S))$. Then, (A18) can be rewritten as

$$\begin{aligned} \sigma_S^2 &= \sum_{m^S=0}^{\infty} \sum_{n=0}^{\infty} \sum_{\hat{n}=0, \hat{n}>0}^{\infty} \sum_{\ell_{1:n} \in \mathbb{L}_n} \sum_{k=0}^{\min(n, \hat{n})} \sum_{\{\ell'_{1:k}\} \in \{\ell_{1:n}\}} \frac{\Omega_{n,m^S}}{m^S!n!} \cdot \frac{1}{n+\hat{n}-k} \int_{\mathbb{Z}_{\hat{n},m^S}^S} \int_{\mathbb{X}_n} q(\mathbf{X}_n, Z_{m^S}^S) \\ &\cdot \left(\sum_{\ell' \in \{\ell'_{1:k}\}} \min(c^2, \|x_{\ell'} - \hat{x}_{\ell'}(Z_{m^S}^S)\|^2) + c^2(n+\hat{n}-2k) \right) dx_{1:n} dz_{1:m^S}^S \\ &= \sum_{m^S=0}^{\infty} \sum_{n=0}^{\infty} \sum_{\hat{n}=0, \hat{n}>0}^{\infty} \sum_{k=0}^{\min(n, \hat{n})} \sum_{\ell'_{1:k} \in \mathbb{L}_k} \frac{\Omega_{n,m^S} \Psi_{\hat{n},n,m^S}}{m^S!k!(n-k)!} \cdot \frac{1}{n+\hat{n}-k} \\ &\cdot \left(\sum_{\ell' \in \{\ell'_{1:k}\}} \min(c^2, \frac{1}{\Psi_{\hat{n},n,m^S}} \sum_{(\ell_{1:n}-\ell') \in \mathbb{L}_{n-1}} \int_{\mathbb{Z}_{\hat{n},m^S}^S} \int_{\mathbb{X}_n} q(\mathbf{X}_n, Z_{m^S}^S) \|x_{\ell'} - \hat{x}_{\ell'}(Z_{m^S}^S)\|^2 dx_{1:n} dz_{1:m^S}^S) + c^2(n+\hat{n}-2k) \right) \end{aligned} \tag{A19}$$

where $\Psi_{\hat{n},n,m^S}$ is defined in (34) and finally obtained as (39), $\ell_{1:n} - \ell'$ denotes the residual label sequence of $\ell_{1:n}$ after the label ℓ' is separately excluded from $\ell_{1:n}$.

Note that the estimate $\hat{x}_{\ell'}(Z_{m^S}^S)$ involved in (A19) is independent of \mathbf{X}_n . So, the integral term in the last line of (A19) is obtained as

$$\begin{aligned} & \sum_{(\ell_{1:n}-\ell') \in \mathbb{L}_{n-1}} \int_{\mathbb{Z}_{\hat{n},m^S}^S} \int_{\mathbb{X}_n} q(\mathbf{X}_n, Z_{m^S}^S) \|x_{\ell'} - \hat{x}_{\ell'}(Z_{m^S}^S)\|^2 dx_{1:n} dz_{1:m^S}^S \\ &= \int_{\mathbb{Z}_{\hat{n},m^S}^S} \int_{\mathbb{X}_1} \left[\sum_{(\ell_{1:n}-\ell') \in \mathbb{L}_{n-1}} \int_{\mathbb{X}_{n-1}} q(\mathbf{X}_n, Z_{m^S}^S) d(x_{1:n} - x_{\ell'}) \right] \|x_{\ell'} - \hat{x}_{\ell'}(Z_{m^S}^S)\|^2 dx_{\ell'} dz_{1:m^S}^S \quad (\text{A20}) \\ &= \int_{\mathbb{Z}_{\hat{n},m^S}^S} \int_{\mathbb{X}_1} q_n(x, \ell', Z_{m^S}^S) \sum_{l=1}^L (x_{\ell'}^l - \hat{x}_{\ell'}^l(Z_{m^S}^S))^2 dx_{\ell'} dz_{1:m^S}^S \end{aligned}$$

where the last line is obtained according to the marginal density $q_n(x, Z_{m^S}^S)$ of (25) and the definition of 2-norm, L is the dimension of x , $\int_{\mathbb{X}_{n-1}} \cdot d(x_{1:n} - x_{\ell'})$ denotes the residual integral of $\int_{\mathbb{X}_n} \cdot dx_{1:n}$ after the term $\int_{\mathbb{X}_1} \cdot dx_{\ell'}$ is separately excluded from $\int_{\mathbb{X}_n} \cdot dx_{1:n}$.

Since **A2** has shown that the estimator $\hat{x}_{\ell'}(Z_{m^S}^S)$ is unbiased, the information inequality of (7) can be applied to (A20),

$$\int_{\mathbb{Z}_{\hat{n},m^S}^S} \int_{\mathbb{X}_1} q_n(x, \ell', Z_{m^S}^S) (x_{\ell'}^l - \hat{x}_{\ell'}^l(Z_{m^S}^S))^2 dx_{\ell'} dz_{1:m^S}^S \geq [J_{\hat{n},n,m^S}^{-1}(\ell')]^{l,l} \quad l = 1, \dots, L \quad (\text{A21})$$

where the FIM $J_{\hat{n},n,m^S}(\ell')$ is shown in (42). (A21) holds with equality if and only if $q_n(x, \ell', Z_{m^S}^S)$ satisfies the distribution of exponential family.

Substituting (A21) into (A20) and then (A19), we get

$$\begin{aligned} \sigma_S^2 &\geq \sum_{m^S=0}^{\infty} \sum_{n=0}^{\infty} \sum_{\hat{n}=0, n+\hat{n}>0}^{\infty} \sum_{k=0}^{\min(n,\hat{n})} \sum_{\ell'_{1:k} \in \mathbb{L}_k} \frac{\Omega_{n,m^S} \Psi_{\hat{n},n,m^S}}{m^S! k! (n-k)!} \cdot \frac{1}{n+\hat{n}-k} \left(\sum_{\ell' \in \{\ell'_{1:k}\}} \min\left(c^2, \frac{1}{\Psi_{\hat{n},n,m^S}} \sum_{l=1}^L [J_{\hat{n},n,m^S}^{-1}(\ell')]^{l,l}\right) + c^2(n+\hat{n}-2k) \right) \\ &= \sum_{m^S=0}^{\infty} \sum_{n=0}^{\infty} \sum_{\hat{n}=0, n+\hat{n}>0}^{\infty} \sum_{k=0}^{\min(n,\hat{n})} \sum_{\ell' \in \mathbb{L}_1} \frac{\Omega_{n,m^S} \Psi_{\hat{n},n,m^S}}{m^S! (n-k)!} \cdot \frac{1}{n+\hat{n}-k} \left(k \cdot \min\left(c^2, \frac{1}{\Psi_{\hat{n},n,m^S}} \sum_{l=1}^L [J_{\hat{n},n,m^S}^{-1}(\ell')]^{l,l}\right) + c^2(n+\hat{n}-2k) \right) \end{aligned} \quad (\text{A22})$$

where the second line is derived from the fact that the FIM $J_{\hat{n},n,m^S}(\ell')$ is the same for all $\ell' \in \{\ell'_{1:k}\}$ given $\ell'_{1:k} \in \mathbb{L}_k$.

Finally, (40) can be obtained by substituting the notation $\varepsilon_{k,\hat{n},n}$ defined in (41) into (A22). \square .

References

- Blackman, S.; Popoli, R. *Design and Analysis of Modern Tracking Systems*; Artech House: Norwood, MA, USA, 1999.
- Hero, A.O.; Castanón, D.A.; Cochran, D.; Kastella, K. *Foundations and Applications of Sensor Management*; Springer: New York, NY, USA, 2008.
- Tharmarasa, R.; Kirubarajan, T.; Hernandez, M.L.; Sinha, A. PCRLB-based multisensor array management for multitarget tracking. *IEEE Trans. Aerosp. Electron. Syst.* **2007**, *43*, 539–555. [[CrossRef](#)]
- Tharmarasa, R.; Kirubarajan, T.; Sinha, A.; Lang, T. Decentralized sensor selection for large-scale multisensor-multitarget tracking. *IEEE Trans. Aerosp. Electron. Syst.* **2011**, *47*, 1307–1324. [[CrossRef](#)]
- Fu, Y.; Ling, Q.; Tian, Z. Distributed sensor allocation for multi-target tracking in wireless sensor networks. *IEEE Trans. Aerosp. Electron. Syst.* **2012**, *48*, 3538–3553. [[CrossRef](#)]
- Mohammadi, A.; Asif, A. Decentralized computation of the conditional posterior Cramér-Rao lower bound: Application to adaptive sensor selection. In Proceedings of the 38th IEEE International Conference on Acoustics, Speech and Signal Processing, Vancouver, BC, Canada, 26–31 May 2013.
- Herath, S.C.K.; Pathirana, P.N. Optimal sensor arrangements in angle of arrival (AoA) and range based localization with linear sensor arrays. *Sensors* **2013**, *13*, 12277–12294. [[CrossRef](#)] [[PubMed](#)]
- Wang, Z.; Shen, X.; Wang, P.; Zhu, Y. The Cramér-Rao bounds and sensor selection for nonlinear systems with uncertain observations. *Sensors* **2018**, *18*, 1103. [[CrossRef](#)] [[PubMed](#)]

9. Mahler, R. *Advances in Statistical Multisource-Multitarget Information Fusion*; Artech House: Norwood, MA, USA, 2014.
10. Vo, B.N.; Singh, S.; Doucet, A. Sequential Monte Carlo methods for multi-target filtering with random finite sets. *IEEE Trans. Aerosp. Electron. Syst.* **2005**, *41*, 1224–1245.
11. Vo, B.N.; Ma, W.K. The Gaussian mixture probability hypothesis density filter. *IEEE Trans. Signal Process.* **2006**, *54*, 4091–4104. [[CrossRef](#)]
12. Zhang, Q.; Song, T.L. Improved bearings-only multi-target tracking with GM-PHD filtering. *Sensors* **2016**, *16*, 1469. [[CrossRef](#)] [[PubMed](#)]
13. Vo, B.T.; Vo, B.N.; Cantoni, A. Analytic implementations of the cardinalized probability hypothesis density filter. *IEEE Trans. Signal Process.* **2007**, *55*, 3553–3567. [[CrossRef](#)]
14. Si, W.; Wang, L.; Qu, Z. Multi-target tracking using an improved Gaussian mixture CPHD filter. *Sensors* **2016**, *16*, 1964. [[CrossRef](#)] [[PubMed](#)]
15. Vo, B.T.; Vo, B.N.; Cantoni, A. The cardinality balanced multi-target multi-Bernoulli filter and its implementations. *IEEE Trans. Signal Process.* **2009**, *57*, 409–423.
16. He, X.; Liu, G. Cardinality balanced multi-target multi-Bernoulli filter with error compensation. *Sensors* **2016**, *16*, 1399. [[CrossRef](#)] [[PubMed](#)]
17. Vo, B.T.; Vo, B.N. Labeled random finite sets and multi-object conjugate priors. *IEEE Trans. Signal Process.* **2013**, *61*, 3460–3475. [[CrossRef](#)]
18. Vo, B.N.; Vo, B.T.; Phung, D. Labeled random finite sets and the Bayes multi-target tracking filter. *IEEE Trans. Signal Process.* **2014**, *62*, 6554–6567. [[CrossRef](#)]
19. Liu, C.; Sun, J.; Lei, P.; Qi, Y. δ -generalized labeled multi-Bernoulli filter using amplitude information of neighboring cells. *Sensors* **2018**, *18*, 1153.
20. Reuter, S.; Vo, B.T.; Vo, B.N.; Dietmayer, K. The labeled multi-Bernoulli filter. *IEEE Trans. Signal Process.* **2014**, *62*, 3246–3260.
21. Papi, F.; Vo, B.N.; Vo, B.T.; Fantacci, C.; Beard, M. Generalized labeled multi-Bernoulli approximation of multi-object densities. *IEEE Trans. Signal Process.* **2015**, *63*, 5487–5497. [[CrossRef](#)]
22. Fantacci, C.; Papi, F. Scalable multisensor multitarget tracking using the marginalized δ -GLMB density. *IEEE Signal Process Lett.* **2016**, *23*, 863–867. [[CrossRef](#)]
23. Fantacci, C.; Vo, B.N.; Vo, B.T.; Battistelli, G.; Chisci, L. Robust fusion for multisensor multiobject tracking. *IEEE Signal Process Lett.* **2018**, *25*, 640–644. [[CrossRef](#)]
24. Williams, J.L. Marginal multi-Bernoulli filters: RFS derivation of MHT, JIPDA and association-based MeMBer. *IEEE Trans. Aerosp. Electron. Syst.* **2015**, *51*, 1664–1687. [[CrossRef](#)]
25. García-Fernández Á, F.; Williams, J.L.; Granström, K.; Svensson, L. Poisson multi-Bernoulli mixture filter: Direct derivation and implementation. *IEEE Trans. Aerosp. Electron. Syst.* **2018**, *54*, 1883–1901. [[CrossRef](#)]
26. Granström, K.; Svensson, L.; Xia, Y.; Williams, J.; García-Fernández Á, F. Poisson multi-Bernoulli mixture trackers: Continuity through random finite sets of trajectories. In Proceedings of the 21st International Conference on Information Fusion, Cambridge, UK, 10–13 July 2018.
27. Ristic, B.; Vo, B.N. Sensor control for multi-object state-space estimation using random finite sets. *Automatica* **2010**, *46*, 1812–1818. [[CrossRef](#)]
28. Ristic, B.; Vo, B.N.; Clark, D. A note on the reward function for PHD filters with sensor control. *IEEE Trans. Aerosp. Electron. Syst.* **2011**, *47*, 1521–1529. [[CrossRef](#)]
29. Hoang, H.G.; Vo, B.T. Sensor management for multi-target tracking via multi-Bernoulli filtering. *Automatica* **2014**, *50*, 1135–1142. [[CrossRef](#)]
30. Gostar, A.K.; Hoseinnezhad, R.; Bab-Hadiashar, A. Multi-Bernoulli sensor control via minimization of expected estimation errors. *IEEE Trans. Aerosp. Electron. Syst.* **2015**, *51*, 1762–1773. [[CrossRef](#)]
31. Beard, M.; Vo, B.T.; Vo, B.N.; Arulampalam, S. Sensor control for multi-target tracking using Cauchy-Schwarz divergence. In Proceedings of the 18th International Conference on Information Fusion, Washington, DC, USA, 6–9 July 2015.
32. Jiang, M.; Yi, W.; Kong, L. Multi-sensor control for multi-target tracking using Cauchy-Schwarz divergence. In Proceedings of the 19th International Conference on Information Fusion, Heidelberg, Germany, 5–8 July 2016.
33. Gostar, A.K.; Hoseinnezhad, R.; Rathnayake, T.; Wang, X.; Bab-Hadiashar, A. Constrained sensor control for labeled multi-Bernoulli filter using Cauchy-Schwarz divergence. *IEEE Signal Process Lett.* **2017**, *24*, 1313–1317. [[CrossRef](#)]

34. Wang, X.; Hoseinnezhad, R.; Gostar, A.K.; Rathnayake, T.; Xu, B.; Bab-Hadiashar, A. Multi-sensor control for multi-object Bayes filters. *Signal Process.* **2018**, *14*, 260–270. [[CrossRef](#)]
35. Rezaeian, M.; Vo, B.N. Error bounds for joint detection and estimation of a single object with random finite set observation. *IEEE Trans. Signal Process.* **2010**, *58*, 1493–1506. [[CrossRef](#)]
36. Arulampalam, S.; Maskell, S.; Gordon, N.J.; Clapp, T. A tutorial on particle filters for on-line non-linear/non-Gaussian Bayesian tracking. *IEEE Trans. Signal Process.* **2002**, *50*, 174–188. [[CrossRef](#)]
37. Wright, S.J. Coordinate descent algorithms. *Math. Program.* **2015**, *151*, 3–34. [[CrossRef](#)]
38. Hristakeva, M.; Shrestha, D. Solving the 0-1 knapsack problem with genetic algorithms. In Proceedings of the 37th Midwest Instruction and Computing Symposium, Morris, MN, USA, 16–17 April 2004.
39. Schuhmacher, D.; Vo, B.T.; Vo, B.N. A consistent metric for performance evaluation of multi-object filters. *IEEE Trans. Signal Process.* **2008**, *56*, 3447–3457. [[CrossRef](#)]
40. Ristic, B.; Vo, B.N.; Clark, D.; Vo, B.T. A metric for performance evaluation of multi-target tracking algorithms. *IEEE Trans. Signal Process.* **2011**, *59*, 3452–3457. [[CrossRef](#)]
41. Rahmathullah, A.S.; Garcia-Fernandez, A.F.; Svensson, L. Generalized optimal sub-pattern assignment metric. In Proceedings of the 20th International Conference on Information Fusion, Xi'an, China, 10–13 July 2017.
42. Zhou, B.; Bose, N.K. Multitarget tracking in clutter: Fast algorithms for data association. *IEEE Trans. Aerosp. Electron. Syst.* **1993**, *29*, 352–363. [[CrossRef](#)]
43. Battistelli, G.; Chisci, L.; Fantacci, C.; Farina, A.; Graziano, A. Consensus CPHD filter for distributed multitarget tracking. *IEEE J. Sel. Top. Sign. Proces.* **2013**, *7*, 508–520. [[CrossRef](#)]
44. Battistelli, G.; Chisci, L. Kullback-Leibler average, consensus on probability densities and distributed state estimation with guaranteed stability. *Automatica* **2014**, *50*, 707–718. [[CrossRef](#)]
45. Davis, P.J.; Rabinowitz, P.; Rheinboldt, W. *Methods of Numerical Integration*, 2nd ed.; Mineola; Dover Publications: New York, NY, USA, 2007.
46. Anderson, B.D.; Moore, J.B. *Optimal Filtering*; Prentice-Hall: Englewood Cliffs, NJ, USA, 1979.
47. Julier, S.J.; Uhlmann, J.K. Unscented filtering and nonlinear estimation. *Proc. IEEE* **2004**, *92*, 401–422. [[CrossRef](#)]
48. Luenberger, D.G.; Ye, Y. *Linear and Nonlinear Programming*, 4th ed.; Springer: New York, NY, USA, 2015.
49. Li, X.R.; Jilkov, V.P. Survey of maneuvering target tracking, part V: Multiple-model methods. *IEEE Trans. Aerosp. Electron. Syst.* **2005**, *41*, 1255–1321.
50. Vo, B.N.; Vo, B.T.; Hoang, H.G. An efficient implementation of the generalized labeled multi-Bernoulli filter. *IEEE Trans. Signal Process.* **2017**, *65*, 1975–1987. [[CrossRef](#)]
51. Beard, M.; Vo, B.T.; Vo, B.N. A solution for large-scale multi-object tracking. *arXiv*. 2018. Available online: <https://arxiv.org/abs/1804.06622> (accessed on 22 November 2018).

